

Kompendium

Forbedringsarbeid og Statistisk Prosesskontroll (SPC)

Hovedkilder:

Raymond G. Carey: "Measuring Quality Improvement in Healthcare. A Guide to Statistical Process Control Applications"

Raymond G. Carey: "Improving healthcare with control charts. Basic and advanced SPC methods and case studies"

Andre kilder:

Dag Hofoss: "Dokumentasjon av forbedringer ved hjelp av statistisk prosesskontroll". Den norske lægeforening.

Dag Hofoss: "Statistisk prosesskontroll – en enkel måte å undersøke om det går bedre eller verre på". inFORUM november 2000

Adeleide Skard Brandrud og Lill De Greve: "Erfaringer med bruk av tidsserieanalyser". inFORUM juni 2001

Kvalitetsforbedring i psykisk helsevern: Prosessforbedring i klinisk virksomhet. Oslo, Helsetilsynet, 2001. 63s. IK-2756.

Diverse andre artikler av ulike forfattere

Stiftelsen GRUK

Gruppe for kvalitetsutvikling

i sosial- og helsetjenesten

Nedre Hjellegt. 1

3724 SKIEN

Tlf 35 58 40 40

E-mail: post@gruk.no

Kontaktperson: Bjørnar Nyen

Innhold

1	LITT OM FORBEDRINGPROSESSER	5
1.1	Generelt	5
1.2	Vanlig, viktig, vanskelig, virkelighetsnært.....	5
1.3	Demings sirkel.....	5
1.4	Flytskjema over prosessforbedringsarbeid	6
1.5	Om å sette mål.....	7
1.6	Prosesshierarki.....	7
1.7	Brukermedvirkning.....	8
1.8	Om målinger før og etter	8
2	OM Å FORSTÅ VARIASJON.....	8
2.1	Naturlig og spesiell variasjon.....	8
2.2	Konsekvenser av ikke å forstå variasjon	9
3	OM FORBEREDELSE TIL DATASAMLING	9
3.1	Identifisering av en mulighet for forbedring.....	9
3.2	Prioritering av muligheter for forbedring.....	10
3.3	Organisering av team	10
3.4	Klargjøring av prosessen med flytskjemaer.....	10
3.5	Standardisering av prosessen	11
3.6	Definering av viktig kvalitetsegenskap (Key Quality Characteristics).....	11
3.7	Å utvikle operasjonelle definisjoner.....	11
4	DATASAMLING	11
4.1	Data versus informasjon	11
4.2	Momenter til en datasamlingsplan.....	12
4.2.1	Hvorfor vil man samle disse dataene?	12
4.2.2	Hvilket nøkkelmål for kvalitet vil man samle data til?.....	13
4.2.3	Hvor ofte og hvor lenge vil man samle data?	13
4.2.4	Vil man benytte utvalg av data? I tilfelle – hva slags utvalgsdesign skal man velge?	13
4.2.5	Skal man kjøre en pilotstudie før man starter å samle data i stor skala?	14
4.2.6	Hvordan skal man samle dataene?.....	14
4.2.7	Vil datainnsamlingen ha negative følger for noen – pasienter, stab eller pårørende?.....	15
4.2.8	Hvordan skal dataene kodes, redigeres, verifiseres og analyseres?.....	15
4.2.9	Hvordan skal dataene brukes og hvem skal ha tilgang til rådataene og resultatene?	15

5	DIAGRAMMER.....	15
5.1	Run-diagram	15
5.1.1	Hva er et run-diagram ?	16
5.1.2	Å lage et run-diagram	16
5.1.3	Hva er et run ?.....	16
5.1.4	Hva er testene for spesiell variasjon ("special cause") ?	16
5.2	Kontroll-diagrammer	17
5.2.1	Grunnleggende elementer	18
5.2.2	Hvordan oppdage spesiell variasjon	18
5.2.3	Finnes det situasjoner der testene for spesiell variasjon bør modifiseres?	19
5.2.4	Hvor mange datapunkter kreves for et kontroll-diagram?.....	19
5.3	Å velge passende kontrolldiagram	19
5.3.1	Identifisere type data.....	19
5.3.2	Om å bestemme hvilket diagram som passer best	20
5.3.3	Om subgrupper	20
5.4	Ulike kontrolldiagrammer	20
5.4.1	P-diagram.....	20
5.4.2	U-diagram	21
5.4.3	C-diagram	22
5.4.4	I-diagram (= XMR-diagram)	22
5.4.5	X-bar og R-diagram	22
5.4.6	X-bar og S-diagram	22
5.5	Hvilket kontroll-diagram er best?	23
5.6	Subgruppestørrelse.....	23
5.6.1	Subgruppestørrelse for p-diagrammer	23
5.6.2	Subgruppestørrelse for c-diagrammer og u-diagrammer.....	23
6	APPENDIX:	24
6.1	Diverse begreper	24
6.2	Litteratur	25
6.2.1	Bøker/tidsskrift	25
6.2.2	Dataprogrammer	25
7	EKSEMPELSAMLING.....	26
7.1	Run-diagram	26
7.2	I-diagram	26
7.3	C-diagram.....	28
7.4	Ubesvarte telefoner – I-diagram og P-diagram	28
7.5	Ekstra mye å gjøre på legevakta – I-diagram	31
7.6	Ventetid på legevakt X-bar S-diagram	32
7.7	Ventetid på telefon – X-bar R-diagram og X-bar S-diagram.....	33
7.8	INR og Marevan-dosering	34
7.9	Mer om spesiell variasjon	37

SPC og forbedringsarbeid i helse- og sosialtjenesten

Det skjer stadige endringer av prosesser i helse- og sosialtjenesten. De fleste planlagte endringene skjer for å bedre resultater i en eller flere av dimensjonene i arbeidet: faglig standard, ressursutnyttelse, brukertilfredshet, tilgjengelighet, risiko for skader og bivirkninger, rettfredig fordeling.

Ofte vil det være ønskelig å måle hva slags effekt endringen av prosessene skaper. For det første er det viktig for helsearbeiderne selv å vite om anstrengelsene i forbedringsarbeidet gir effekt. Men det er også viktig å måle for å dekke kravet til dokumentasjon av kvalitet; et krav som i økende grad kommer både fra brukere, myndigheter og medier. Det er også fokus på muligheter for innsparinger.

Det er noen problemer med målinger. Uten kunnskap i statistikk, blir konklusjonene lett feil og ofte direkte villedende. I den statistikken man innen medisinen er mest kjent med, trengs ofte store datamengder og er ofte vanskelig å bruke i det daglige arbeidet.

Vi trenger et statistiske verktøy for målingene som er dynamisk, som ikke krever store datamengder for å gi sikre resultater, som ikke er for komplisert og tidkrevende å bruke og som ikke er for vanskelig å forstå.

Tidsserieanalyser har i mange år hatt en stor plass innen mange andre felt som industri og økonomi. I løpet av de siste 20 årene har dette i økende grad blitt brukt også inne helsetjenesten i mange land, særlig i form av den typen for tidsserieanalyse som kalles **statistisk prosesskontroll** (SPC¹). Flere og flere oppdager nå at dette er et nyttig verktøy for helsetjenesten. Noen få har brukt det gjennom noen år i Norge, blant andre Aleidis Skard Branderud, Lill De Greve, Ove Kjell Andersen, Dag Hofoss, Otto Brun Pedersen.

Mange har hentet kunnskap og inspirasjon fra Raymond G. Carey, som har skrevet lærebøker i SPC og som holdt et to dagers kurs i Norge i mars 2003.

Carey skriver: "Statistisk prosesskontroll (SPC) er en filosofi, en strategi og et sett med metoder for vedvarende forbedring av systemer, prosesser og resultater. SPC-tilnærmingen er basert på å lære av data og har sitt grunnlag i variasjonsteori (forståelse av naturlige og spesielle variasjoner). SPC-strategien innbefatter konseptene om analytisk studie, prosess-tenkning, forebygging, stratifisering, stabilitet, kapabilitet og forutsigbarhet. SPC innbefatter målinger, datasamlings-metoder og planlagt eksperimentering. Grafiske metoder, som Shewhart-diagrammer (vanligvis kalt kontroll-diagrammer), run-diagram, "frequency plots", histogrammer, Pareto-analyse, scatter plott og flytskjemaer er de primære verktøyene brukt i SPC."

Dette høres komplisert ut, men man kommer et godt stykke på vei ved å ta i bruk bare deler av dette, så kan man heller utvide repertoaret etter hvert.

Dette kompendiet er i første rekke skrevet for de helse- og sosialarbeidere som er i prosjekter hvor GRUK er involvert. Det er ikke et "ferdig" produkt – kommentarer mottas med takk! Hele eller deler av kompendiet kan printes ut til personlig bruk, men det er ikke tillatt å anvende det til undervisning, utgivelse eller distribusjon uten etter avtale med GRUK.

Skien, februar 2004

Bjørnar Nyen

¹ Vi bruker den engelske forkortelsen SPC som står for "Statistical Process Control"

1 Litt om forbedringprosesser

1.1 Generelt

Det er viktig å ha plan og struktur i det kontinuerlige forbedringsarbeid på et arbeidssted. Det er noen prinsipper som har vist seg å være av stor betydning i dette:

- Ledelsen må involveres (forbedringsarbeidet forankres ovenfra)
- Medarbeiderne må involveres (forbedringsarbeidet forankres nedenfra)
- Det må være klart hvem som har ansvar for å styre forbedringsarbeidet
- Noen må fristilles for å tilrettelegge/følge opp forbedringsarbeidet (skrivearbeid, gjennomføring og bearbeiding av målinger og intervjuer etc)
- Tverrfaglig deltakelse er nødvendig
- Opplæringen og veiledningen må skje praksisnært

Prosessforbedring dreier seg om å kartlegge, analysere og forbedre sentrale prosesser. Perspektivet bør som regel være pasientens vei gjennom behandlingsapparatet på tvers av profesjoner og tjenestested.

1.2 Vanlig, viktig, vanskelig, virkelighetsnært

Når man skal velge ut prosesser for forbedring bør man prioritere områder som er vanlige og viktige og som har erkjente mangler.

- | | |
|---------------------------|---|
| • Vanlig | Området omfatter mange pasienter |
| • Viktig | Området har betydning for pasientenes helse og livskvalitet |
| • Vanskelig | Området har erkjente mangler, kan være vanskelig å endre |
| • Virkelighetsnært | Det skal være realistisk å få til endring |

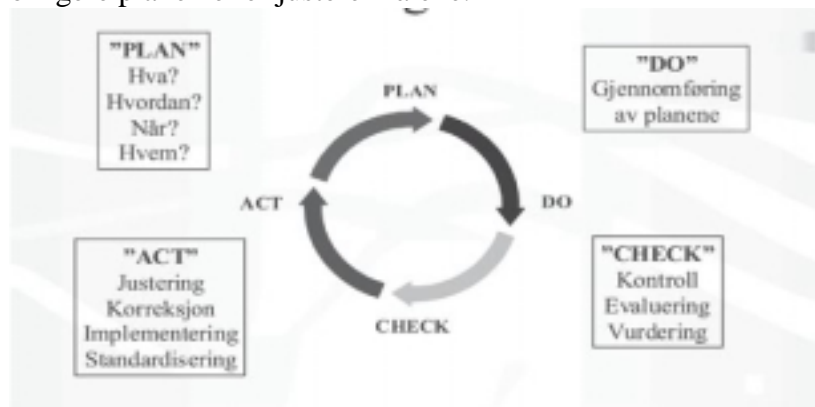
1.3 Demings sirkel

På de fleste arbeidssteder finnes mange områder for mulig forbedring. Man må prioritere blant disse mulighetene, legge en plan for endring og så implementere dem en om gangen og få målt effekten av hver endring – alt i en PDCA syklus (Plan – Do – Check – Act). Denne syklusen kalles også ”Demings sirkel” (figur 1). Den brukes som modell av mange som arbeider med kvalitetsteori og kvalitetsarbeid. Her fokuseres det på betydningen av å lære av erfaring.

Demings sirkel illustrerer fire trinn som vil være et godt fundament for all planlegging og evaluering:

1. Planlegg (**Plan**) en endring. Planlegging innebærer å avklare situasjonen i øyeblikket, fastsette de mål man vil nå og nødvendige tiltak for å nå målene. Til planleggingen hører også å sørge for at deltakerne har kunnskap og at nødvendige ressurser (folk, tid, penger) er tilgjengelig.
2. Utfør (**Do**) endringen i et forsøk, på en mindre testgruppe eller hele gruppen; iverksetting av de planlagte tiltakene.

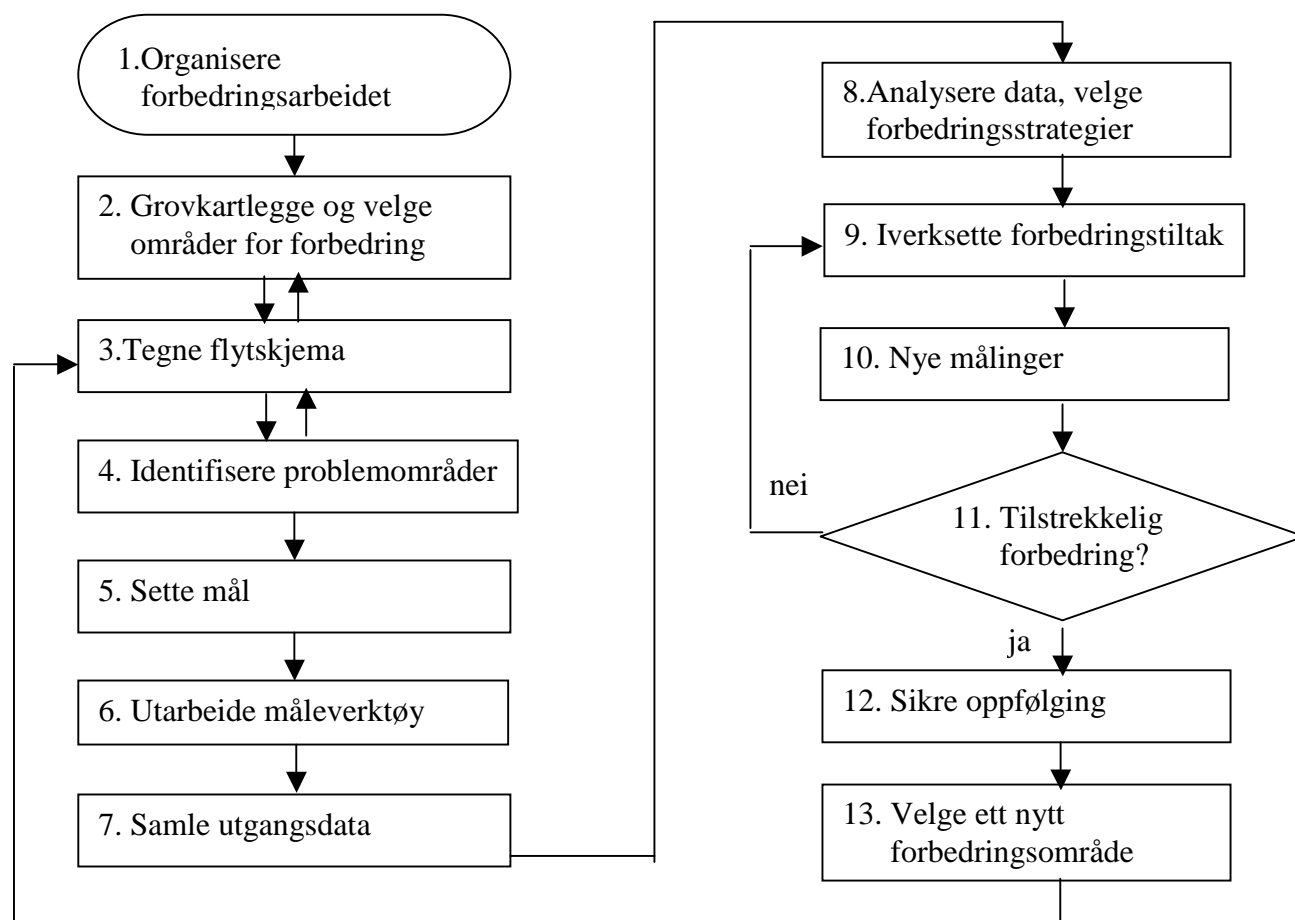
3. Kontroller (**Check**) resultatene. Innsamling og analyse av resultatene i forhold til mål. Har den nye prosessen et nivå av utførelse og/eller naturlig variasjon som er bedre enn i den gamle prosessen?
4. Korrigere (**Act**). Det foretas en oppsummering av prosessen basert på evalueringen i forrige trinn. Dersom man har funnet en god løsning, starter man tiltak som sikrer at den blir innført som rutine og standardisert i hele organisasjonen. Ved avvik må man enten korrigere planen eller justere målene.



Figur 1. Demings sirkel (fra Hofoss).

1.4 Flytskjema over prosessforbedringsarbeid

Innholdet i Demings sirkel kan tegnes som et flytskjema i noe mer detalj (figur 2). Da blir det en beskrivelse av arbeidsgangen i prosessforbedring.



Figur 2. Flytskjema over prosessforbedringsarbeid

De 7 første boksene i dette flytskjemaet dekker innholdsmessig de 7 trinnene Carey har foreslått at man bør ha gått gjennom før man starter datainnsamling i et forbedringsarbeid - se avsnittene 3.1 til 3.7.

Rekkefølgen av boksene i flytskjemaet skal ikke følges slavisk. Særlig for de første 4 punktene kan det være nødvendig å gå litt fram og tilbake for å komme fram til det området man skal ta tak i.

Flytskjemaet har ikke noe definert endepunkt, for forbedringsarbeid er en kontinuerlig prosess, som dreier seg om å utvikle og vedlikeholde kvaliteten i virksomheten.

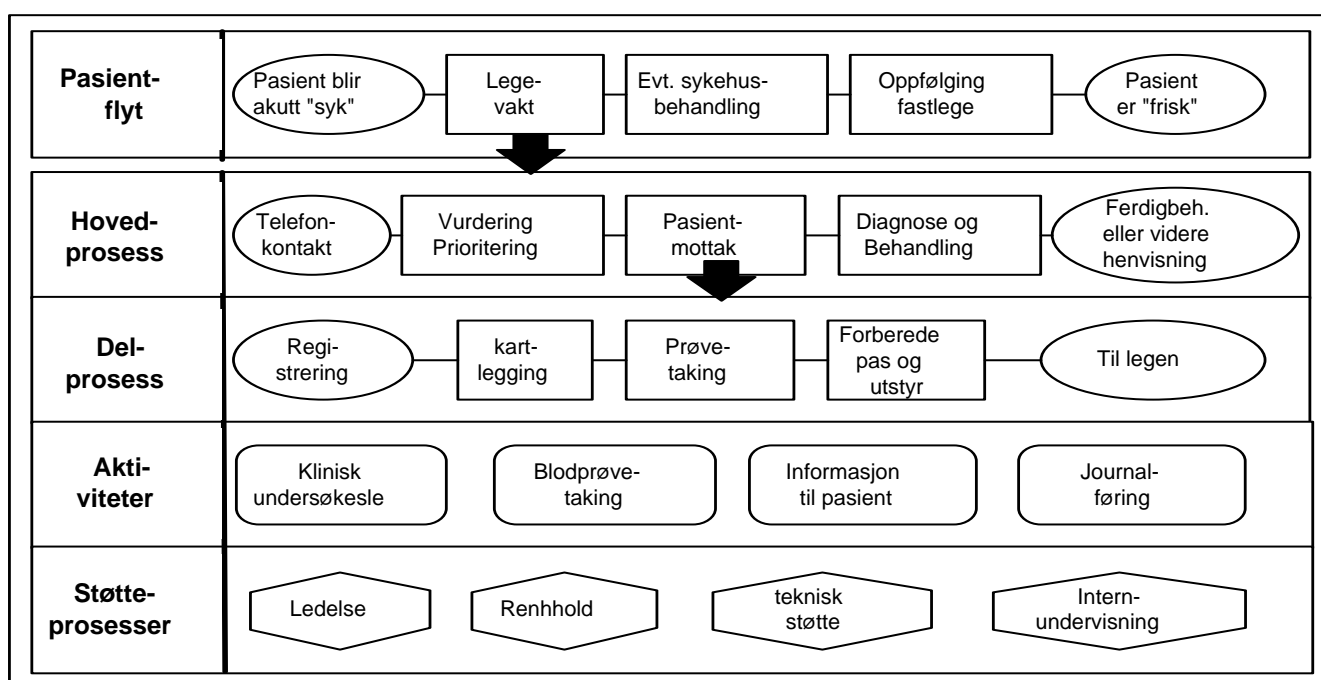
1.5 Om å sette mål

All erfaring tilsier at det er fruktbart å sette mål for hva man vil oppnå ved forbedringsarbeidet. Målsettingen bør være konkret. Den bør være målbar, kvantitativt og/eller kvalitativt. Det er viktig å være realistisk og vurdere om målene som settes er oppnåelige i forhold til ressursene. Man bør også angi når resultatene skal være oppnådd.

1.6 Prosesshierarki

Ofte kan det være nyttig å tenke seg et prosesshierarki med ulike nivåer, jfr figur 3:

- Tegning av pasientflyten hjelper oss til å holde på pasientperspektivet.
- Tegning av hovedprosessen hjelper oss til å beholde oversikten.
- Tegning av delprosesser hjelper oss til å være tilstrekkelige konkrete for å gjøre nødvendig forbedringsarbeid.



Figur 3. Prosesshierarkiet

1.7 Brukermedvirkning

I forbedringsarbeid i helse- og sosialtjenesten er det alltid viktig å ha et brukerfokus. Det viser seg oftest veldig nyttig å ha med en eller flere brukere/pasienter i planleggingen og eventuelt også i gjennomføringen av et forbedringsarbeid.

Dette kan suppleres med, og noen ganger erstattes av, brukerundersøkelser.

1.8 Om målinger før og etter

Representative utgangsdata er nødvendige både for å identifisere kvalitetsproblemer og for å dokumentere forbedringer. Man skal være oppmerksom på at straks man setter fokus på en arbeidsprosess, påvirkes den. Derfor bør man starte målingene så snart som mulig. Innsamling av utgangsdata må være avsluttet før man begynner endringer i arbeidsprosessen. Dette er imidlertid ikke alltid mulig fordi man ønsker forbedring før man rekker å samle utgangsdata.

Resultatene må vurderes i forhold til den målsettingen man hadde i utgangspunktet. Viser målingene at det ikke har skjedd noe betydelig forbedring, må man vurdere om tiltaket ble godt nok gjennomført eller om det er behov for andre typer tiltak.

2 Om å forstå variasjon

Det er to måter å vise variasjon på: statisk visning og dynamisk visning. Det vanligste måten har vært statisk visning presentert i tabeller, histogrammer etc med utregning av gjennomsnitt og med mål for spredning (for eksempel standard avvik, variasjonskoeffisient etc).

Dynamisk visning er å presentere data i et run- eller kontroll-diagram. Det tillater oss å se hvordan dataene varierer over tid.

Hver av disse tilnærmingmåtene har sin plass, og utfordringen er å vite når man bør bruke det ene og når man bør bruke det andre. Som regel er den dynamiske visningen mer nyttig enn den statiske når man skal bruke dataene i forbedringsarbeid.

2.1 Naturlig og spesiell variasjon

Shewart og Deming utviklet forståelsen av disse to typene variasjon.

Naturlig variasjon (common cause variation) er en iboende egenskap ved en prosess – den er tilfeldig og har naturlige, vanlige årsaker. Den har altså ikke en spesiell årsak, og den er uttrykk for prosessens rytme. Shewart kalte den “assignable variation”. Man kan, innen visse grenser, forutsi utfallene i en prosess som bare har naturlig variasjon. Prosessen er stabil og resultatene av prosessen er forutsigbar.

Spesiell variasjon (special cause variation) skyldes irregulære, unaturlige årsaker som ikke er iboende i prosessen. Shewart kalte den “unassignable variation”. Om en prosess har spesiell variasjon, sier vi at den er “ute av kontroll”; den er ustabil og vil være uforutsigbar.

Shewart og Deming forsøkte å understreke følgende viktige punkter:

1. Variasjon forekommer i alt det vi gjør
2. Prosesser som bare har naturlig variasjon er forutsigelige innenfor statistiske grenser
3. Spesiell variasjon bør identifiseres og bli ryddet av veien
4. Å forsøke å forbedre en prosess som har spesiell variasjon vil kunne øke variasjonen og kaste bort ressurser
5. Når spesiell variasjon er fjernet, er det formålstjenlig å vurdere å endre prosessen
6. Om en prosess bare har naturlig variasjon, men fungerer på et uakseptabelt nivå, bør man fokusere på å endre prosess-gjennomsnittet i ønsket retning

Shewhart utviklet kontroll-diagrammer for å skille de to typene variasjon. Deming og andre har videreutviklet disse (se kapittel 5 og 6).

2.2 Konsekvenser av ikke å forstå variasjon

Hvis man ikke har en grunnleggende forståelse av variasjon kan man lett se trender der det ikke er trender. Det er vanlig å se en trend i en prosess lenge før den egentlig er der statistisk sett. Se avsnitt 5.4 og 6.3.

Den andre feilen som ofte gjøres ved manglende forståelse for variasjon er at man klandrer (det er ille!) eller gir anerkjennelse (det er ikke så farlig!) til andre for ting de har liten eller ingen kontroll over. Det som kan synes som en forverring eller en forbedring av prosessen kan være uttrykk for naturlig variasjon, ikke for en reell forbedring eller forverring av prosessen. Arbeidsmiljøet kan bli dårligere, medarbeidere blir redde for ”refs” for negative resultater, skuffelsene blir flere fordi det alltid etter en tilsynelatende forbedring kommer en tilsynelatende forverring av resultatene og folk sliter seg lettere ut (økt stress når resultatene blir ”dårligere”). Når man vurderer variasjoner feil, vil man fort lete etter syndebukker når noe ikke går så bra.

Til slutt vil man, ved manglende forståelse av variasjon, ha vansker med å forstå tidligere ytelse, i mindre grad kunne forutsi framtidig ytelse og også ha vansker med å gjøre signifikante forbedringer i prosessene.

3 Om forberedelser til datasamling

Før man starter datainnsamling, anbefaler Carey at man går gjennom 7 forberedende trinn:

3.1 Identifisering av en mulighet for forbedring

Hvilke prosesser skal vi forbedre- og hvilken prosess skal vi begynne med?

I utvelgelsen av område er en av nøklene heller å bruke brukerperspektivet enn behandlers perspektiv på prosessen.

Noen ganger kan det være nyttig å bruke en *fokusgruppe* som er en utmerket måte å lytte til brukerens stemme på. Her inviteres noen utvalgte brukere/pasienter til å sitte ned med tjenesteyterne i en time eller to for å snakke sammen om ting de mener bør forbedres.

En annen metode er å gjøre en brukerundersøkelse. Dette er ofte et nyttig redskap for å identifisere og prioritere forbedringsområder, men bare om de blir utført profesjonelt og med skjønnsomhet.

Personalets erfaringer er gjerne en nødvendig kilde til å identifisere forbedringsområder. Det kan komme fram ved tegning av flytskjema, ved samtaler, ev. som en fokusgruppe eller som intervju.

Man kan også bruke målinger som utgangspunkt, for eksempel ventetider, medisineringsfeil, kostnader etc.

3.2 Prioritering av muligheter for forbedring

Som regel vil man finne flere muligheter for forbedring enn man har tid og ressurser til å gjennomføre. Man må derfor snevre inn mulighetene til det man har ressurser til og som vil gi best effekt i forhold til ressursene som brukes i forbedringstiltakene.

Et hjelpeverktøy i denne utvelgelsen kan være en Pareto-analyse. Det er en metode for å skille ut ”de få vitale fra de mange trivielle”, jfr det vi har skrevet om vanlig, viktig og vanskelig (1.2). Dette fordi det i de fleste tilfeller er mange faktorer som til sammen gir en effekt, men som regel er det noen få faktorer som har den vesentligste effekten på resultatet. Hvis man går inn og forbedrer faktorer som har liten innvirkning på totalresultatet, risikerer man å bruke mye ressurser med liten effekt.

3.3 Organisering av team

Av og til er det ikke nødvendig å involvere andre enn seg selv og eventuelt en medarbeider for å gå inn i en forbedringsmulighet. Men ved mer komplekse prosesser er det som regel nødvendig å involvere flere. Et tverrfaglig team vil som regel lettere se alle forhold ved et problem og også ha mer kunnskap om hvordan prosessen fungerer enn bare en person.

Man bør stille spørsmål som:

- Har man nødvendig mandat og støtte fra ledelsen?
- Hvem skal være med i teamet?
- Hva slags opplæring er nødvendig for teammedlemmene?
- Hvordan skal teamet dele opp arbeidet for å få det gjort?
- Hvem skal skrive referater fra møtene?
- Bør man ha en tilrettelegger for teamet?
- Hvordan skal teamet holdes orientert underveis?

3.4 Klargjøring av prosessen med flytskjemaer

Når en prosess fungerer dårlig, er det ofte slik at det ikke bare er ett trinn i prosessen som er problematisk, men ofte er det flere trinn og/eller overgangene mellom trinnene som fungerer dårlig. Derfor bør den første oppgaven for teamet være å identifisere alle trinnene i prosessen. For å få gjort det, er det ofte nyttig å tegne et flytskjema. Å gjøre det for en prosess vil hjelpe teamet til å se mer av kompleksiteten av prosessen og også bedre se muligheter for forbedringer. Se ellers punkt 1.6.

3.5 Standardisering av prosessen

For å forbedre en prosess, må prosessen i utgangspunktet utføres på en standardisert måte, skriver Carey. På den annen side kan første del av forbedringsarbeidet nettopp være å endre en ustabil, ikke-standardisert prosess til en stabil prosess, for deretter å fortsette forbedringsarbeidet.

Det kan hende at man har ulike rutiner for ulike deler av dagen. Da bør man enten bli enige om å bruke samme rutiner hele døgnet, eller å se på prosessene i ulike deler av døgnet separat.

3.6 Definerer av viktig kvalitetsegenskap (Key Quality Characteristics)

Når man vet hvordan prosessen er, vil neste trinn være å identifisere det aspektet ved prosessen man vil forsøke å forbedre. Dette kalles ofte for viktig kvalitetsegenskap (Key Quality Characteristic = KCQ). For å identifisere det, bør man starte med å identifisere ulike kvalitetsmål for prosessen (for eksempel: kort ventetid, lav pris, bli møtt med vennlighet, få god informasjon), og deretter velge ut det man mener er viktigst sett med brukerperspektiv.

3.7 Å utvikle operasjonelle definisjoner

Viktige kvalitetsegenskaper er abstrakte konsepter som trengs å bli beskrevet i målbare begreper. Å beskrive nøkkelmål på en slik måte er å utvikle en operasjonell definisjon som gir kommuniserbar mening. Definisjonen må være klar og utvetydig. Den må også spesifisere målemetoden.

For eksempel vil "akseptabel ventetid" være en viktig egenskap for kvalitet. Men hva er "akseptabel ventetid"? Her må teamet i hvert tilfelle finne fram til hva man mener med det og definere det (for eksempel en time). Det å definere konkret akseptabel ventetid er nødvendig for å få målt hvor mange som kommer til innenfor målet "akseptabel ventetid".

Når man har vært gjennom disse 7 trinnene, vil man være klar for å lage en plan for datasamlingen.

4 Datasamling

4.1 Data versus informasjon

Data kan noen ganger være fakta og figurer med lite nytte. Men om data blir satt inn i en sammenheng for analyse, kan de bli nyttige for beslutningstaking og forbedring. Data endres derved til å bli nyttig informasjon.

4.2 Momenter til en datasamlingsplan

Samling av data synes å være lett, men kan egentlig være ganske komplisert. Carey sier at datasamling dreier seg like mye om kunst som om vitenskap og læres etter hvert som man får erfaring med å gjøre det.

Enhver god datainnsamling begynner med å utvikle en plan. Planen bør ikke bare innbefatte de tekniske detaljene om hvordan man skal samle inn dataene, men også angi grunnene for datainnsamlingen og hvordan dataene skal brukes.

Videre følger en del spørsmål og korte kommentarer som kan være til hjelp når man skal lage en plan for datainnsamling. Disse spørsmålene bør drøftes før man begynner datainnsamlingen.

4.2.1 Hvorfor vil man samle disse dataene?

Fra et kvalitetsforbedrings-perspektiv bør data samles slik at teamet kan gjøre endringer på bakgrunn av resultatene.

Det kan være mange grunner til å samle data:

- For å finne forbedringsområder
- For å finne de viktigste årsakene til problemer
- For å forstå variasjonen i prosessen
- For å overvåke prosessen over tid
- For å se effekten av en forandring i prosessen
- For å lage et felles referanse-punkt
- For å lage et sikrere grunnlag for å forutse prosessen videre
- Monitorere konsekvenser av nedskjæringer

Hensikten vil styre hva slags data som skal samles og hvordan de samles.

Data skal samles for å utføre statistiske studier som utgjør et grunnlag for endringsarbeid. Deming beskrev to typer av studier som kan hjelpe oss i forståelsen av hvorfor vi samler data.

1) Beskrivende (enumerative) studier blir gjort på en definert populasjon for en gitt tidsperiode og/eller sted og blir gjort for å beskrive resultatene.

Beskrivende studier egner seg for å svare på følgende spørsmål:

- Hva var gjennomsnittlige utgifter for denne gruppe av pasienter?
- Hva var gjennomsnittslengden av oppholdet med standardavvik for alle hoftelddsprotese-operasjoner siste år?
- Hvor stor prosent av innlagte pasienter var kirurgiske sist år?

2) Analytiske studier gjøres på dynamiske prosesser og endringer. De fokuserer like mye på å forutsi framtida som å beskrive fortida, og de søker å hjelpe med å forstå hvorfor resultatene ble som de ble og hva som kan gjøres for å forbedre prosessene som førte til de observerte resultatene.

Eksempler på aktuelle spørsmålene man kan stille i forhold til analytiske studier er:

- Hva kan sies om prosessene som førte til at den aktuelle gruppen ga disse kostnadene?

- Hva kan vi forutsi om lengden på oppholdet for hofteproteseopererte i det kommende året?
- Hva var årsaken til den observerte nedgangen av innlagte kirurgiske pasienter sist år?
- Hvordan vil vi kunne vite om en forbedringsstrategi vil være effektiv?

4.2.2 Hvilket nøkkelmål for kvalitet vil man samle data til?

Hvordan vil nøkkelmålet bli operasjonelt definert?

Å være klar på hvilket område man vil måle, er helt avgjørende. Når et område er definert, må man utvikle en operasjonell definisjon (jfr avsnitt 3.7).

4.2.3 Hvor ofte og hvor lenge vil man samle data?

Skal man samle data for hver pasient i de neste to måneder? Eller skal man trekke ut en tilfeldig pasient hver dag i en måned? Valgene som går på frekvens og varighet vil være avgjørende for kvaliteten på dataene.

Det er viktig av hensyn til ressursbruk ikke å samle mer data enn nødvendig i forhold til formålet. Noen typer av datasamling som for eksempel brukerintervju, tar lang tid, mens andre typer av datasamling for eksempel liggetid, er fort gjort

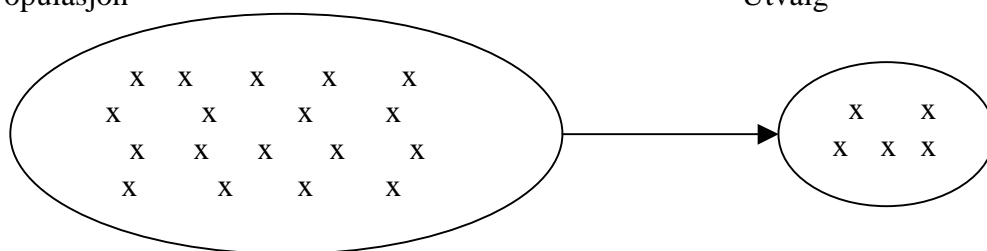
4.2.4 Vil man benytte utvalg av data? I tilfelle – hva slags utvalgsdesign skal man velge?

Her skal vi kort gå gjennom tre utvalgstilnærminger.

Enkelt tilfeldig utvalg ("simple random sampling")

Populasjon

Utvalg



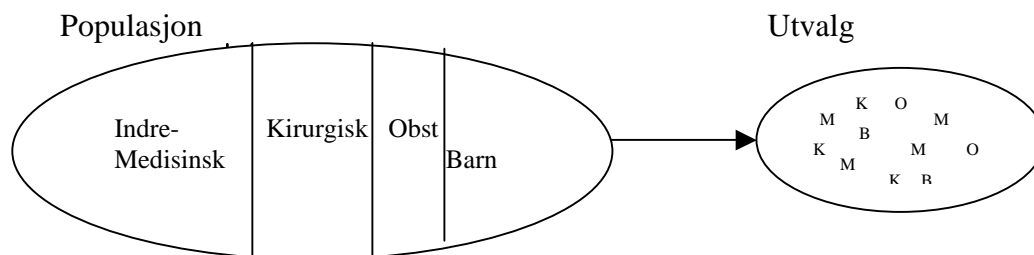
Figur 5: Enkelt tilfeldig utvalg (R. Carey)

Ved tilfeldig utvalg vil hvert datapunkt ha samme sjanse til å bli med i utvalget.

Det er to viktige ulemper ved denne tilnærmingen: 1) den utnytter ikke den kunnskap teamdeltakerne eventuelt har om populasjonen, og 2) den produserer ofte større utvalgs-feil ved samme størrelse på utvalget enn ved stratifisert utvalg.

Enkelt tilfeldig utvalg brukes oftest når man skal gjøre en beskrivende (enumerativ) studie.

Forholdsmessig stratifisert tilfeldig utvalg ("proportional stratified random sampling")



Figur 6: Forholdsmessig stratifisert tilfeldig utvalg (R. Carey)

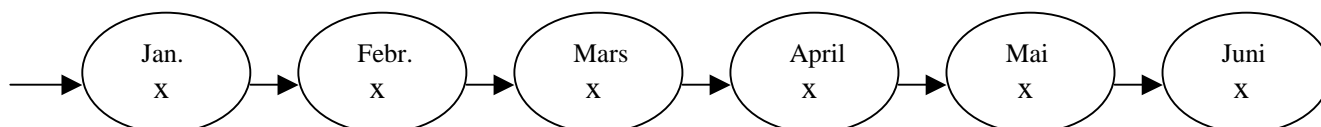
Her deles populasjonen opp i kategorier eller strata. Forholdet mellom de ulike kategoriene bør være omtrent det samme i utvalget som i totalpopulasjonen. Innen hver kategori trekkes det ut data tilfeldig.

Fordelene med denne tilnærmingen er flere. For det første sikrer den representasjon fra de relevante kategoriene i populasjonen. For det andre produserer den mindre variabilitet eller feil enn ved tilfeldig utvalg. For det tredje kan egenskaper av hvert stratum bli undersøkt og analysert for seg.

Men den har også ulemper: Metoden krever nøye kunnskap om populasjonen i hvert stratum. Dessuten vil innsamlingen av data kreve mer ressurser enn ved tilfeldig utvalg.

Rasjonelt utvalg ("judgment sampling").

Dette kalles også ekspert-utvalg ("expert sampling") Deming mente dette var den beste metoden for spesielle typer av analytiske studier, nemlig til \bar{X} -R og \bar{X} -S diagrammer. Rasjonelt utvalg gjøres ved å velge ut en serie av subgrupper fra populasjonen basert på kunnskapen til dem som kjenner prosessen godt. Subgruppene kan trekkes ut enten tilfeldig eller ikke-tilfeldig. Utvalget skjer kontinuerlig, mens det ved de andre utvalgsmetodene skjer ved et fast tidspunkt.



4.2.5 Skal man kjøre en pilotstudie før man starter å samle data i stor skala?

Pilotstudier, også kalt pretester, er en god måte å avsløre hindringer eller vansker for datasamlingen. Man kan tenke seg en pilotstudie som en test på hvor god en datasamlingsplan er. Det gir en anledning til å teste ut ideene og metodene, og til å bestemme om det er nødvendig å gjøre endringer.

4.2.6 Hvordan skal man samle dataene?

Skal man bruke enkelt telleskjema ("data check sheet"), en spørreundersøkelse, fokusgrupper, historiske data, telefonintervjuer, eller en kombinasjon? Det er ofte enklest å bruke historiske

data, men problemet med dem er at de sjelden er samlet for å gi svar på spørsmål om kvalitetsforbedring og er derfor ofte lite egnet.

Et annet spørsmål er hvem som skal ha ansvaret for å samle inn dataene, enten ved å registrere dem, ta dem ut fra datafiler, eller på annen måte. Ofte krever dette mye tid.

Det er også viktig å vurdere hvor mye datainnsamlingen vil koste – hvor mange timer ansatte vil bruke til dette innen vanlig arbeidstid, om det skal betales overtid eller om man må hyre inn og betale noen utenfra

4.2.7 Vil datainnsamlingen ha negative følger for noen – pasienter, stab eller pårørende?

Dette spørsmålet bør vurderes før man setter i gang som en slags ”økologisk sjekk”, gjerne i samarbeid med en etisk komite.

4.2.8 Hvordan skal dataene kodes, redigeres, verifiseres og analyseres?

Ofte blir det store spørsmålet: hvem skal gjøre dette arbeidet? I klinisk forbedringsarbeid har det vist seg at man lykkes best om teamet selv tar ansvar for dataanalysene. Teamet kan hente inn råd fra andre med kunnskap om oppsetting av databaser og statistisk analyse, men teamet bør beholde eierskapet til dataene og analysene.

4.2.9 Hvordan skal dataene brukes og hvem skal ha tilgang til rådataene og resultatene?

Det er viktig å ivareta taushetsplikten i forhold til pasientdata. Teamet må derfor være bevisste på konfidensialitet og datasikkerhet både i planleggingen og gjennomføringen.

5 Diagrammer

5.1 Run-diagram

Shewart utviklet på 1920-tallet et verktøy kalt *run-diagram*.

Run-diagram er veldig enkelt å konstruere og tolke. Det kan brukes på alle typer prosesser og med alle slags data (målte data, telte data, prosenter, forholdstall osv). Et run-diagram kan konstrueres med papir og blyant, man trenger ingen statistisk bearbeiding, og det er enkelt å forstå. Den største svakheten ved run-diagram er at de kan avsløre noen, men ikke alle, spesielle variasjoner. Sensitiviteten er altså dårlig.

Et kontroll-diagram som skal omtales senere, må man gjerne bruke et dataprogram for å få fram, og det er mer komplekst i tolkningen, men har vesentlige fordeler – se punkt 5.2.

5.1.1 Hva er et run-diagram ?

Run-diagram er en måling av en prosess over tid og gir en løpende oversikt over dataene. Tiden er på horisontalaksen. Kvalitetsegenskapen (KQC, se pkt 3.6) er på den vertikale aksen. Senterlinjen i et run-diagram er medianen, mens den i et kontroll-diagram er gjennomsnittet av alle punktene.

5.1.2 Å lage et run-diagram

- Dra en horisontal linje (x-aksen) og merk den med tidsenheten (dager, uker, måneder) eller med sekvensen som dataene er samlet i.
- Dra en vertikal akse (y-aksen) og merk den for å dekke hele spredningen av foreliggende data (\pm ca 20 % for å gi plass for framtidige datapunkter).
- Merk av minimum 15 datapunkter i tidsorden.
- Trekk linje mellom punktene
- Bestem medianen for dataen og trekk en linje for den på kurven.

5.1.3 Hva er et run ?

Et run er definert som ett eller flere etterfølgende datapunkter på samme siden av medianen. Et run kan altså ha ett eller flere punkter. Punkter som ligger på medianen skal ignoreres.

5.1.4 Hva er testene for spesiell variasjon ("special cause") ?

Vanligvis brukes tre tester for å finne spesiell variasjon:

Test nr 1: Spesielt mange eller spesielt få runs.

Når er det så mange eller så få runs at det tilsier at det foreligger en spesiell variasjon? Matematikere har utviklet en tabell som gir svar på dette (tabell 1). Det er ikke nødvendig å skjønnne tabellen!

Det er best å ha minst 16 punkter i et run diagram, eksklusive de på medianen, for å ha tilstrekkelig statistisk styrke til å identifisere en spesiell variasjon. Antall utover 24 punkter gir ikke ytterligere statistisk styrke.

Test 2: Skifte i prosessen

Spesiell variasjon er tilstede hvis et run inneholder spesielt mange punkter. Med 20 eller flere datapunkter, tilsier et run på 8 eller flere at det foreligger spesiell variasjon. Med færre enn 20 punkter, vil et run på 7 også bli ansett å indikere spesiell variasjon.

Antall observasjoner unntatt punkter på medianen	Laveste grense for antall runs	Øvre grense for antall runs
14	4	11
15	4	12
16	5	12
17	5	13
18	6	13
19	6	14
20	6	15
21	7	15
22	7	16
23	8	16
24	8	17
25	9	17
26	9	18
27	9	19
28	10	19
29	10	20
30	10	21

Tabell 1. Sannsynlighetstabell for rundiagram (Fra F.Sweed and C.Eisenhart, "Tables for Testing Randomness of Groupings in a Sequence of Alternative")

Test 3: Trend

En trend er en spesiell variasjon. En trend er definert som en uvanlig lang serie av etterfølgende økende eller synkende verdier. Det er uenighet om antall, men vanligvis settes grensen på minst 6 eller 7 datapunkter. Her regner man med punkter på medianen, men ikke punkter som repeterer foregående verdi.

Eksempel på run-diagram på side 26.

5.2 Kontroll-diagrammer

Både run-diagram og kontroll-diagram har samme formål: å skille naturlig variasjon fra spesiell variasjon i dataene som kommer ut av en prosess. Typen variasjon skal hjelpe til med tilnærmingen for å bedre prosessene og med å evaluere effektivitet av forbedringsanstrengelser.

Kontroll-diagrammer er mer sensitive og mer kraftfulle verktøy enn et run-diagram. I tillegg til en senterlinje (gjennomsnitt) har kontroll-diagram kontrollgrenser som gir ytterligere tester for å identifisere spesielle variasjoner som et run diagram ikke avslører. Kontroll-diagram vil også gi mulighet til å estimere yteevnen (kapabiliteten) av en stabil prosess, dvs å kunne forutsi hvilke grenser den stabile prosessen vil holde seg innenfor.

Den økte sensitiviteten og kraften har en pris. Det er nødvendig å velge riktig type kontroll-diagram for de foreliggende dataene, og det er nødvendig med datamaskin for å konstruere et kontroll-diagram.

5.2.1 Grunnleggende elementer

Senterlinjen er gjennomsnittet av dataverdiene (aritmetisk gjennomsnitt) i et kontroll-diagram. Gjennomsnittet betegnes ofte som \bar{x} (x bar)

Den økte sensitiviteten og kraften i kontroll-diagram ligger i at man bruker *øvre og nedre kontrollgrense*. Ved dette kan man, hvis ingen spesiell variasjon foreligger, estimere den fremtidige kapasiteten av prosessen ut fra gjennomsnittet og kontrollgrensene.

Den øvre og nedre kontrollgrensen faller sammen med ± 3 SD (standardavvik²) fra gjennomsnittet. Med en normalfordelt data vil 99,73 prosent av datapunktene ligge innenfor kontrollgrensene. Det er altså 0,27 % sjanse for at et datapunkt ligger utenfor kontrollgrensene på grunn av tilfeldig variasjon. Med en slik kontrollgrense er sannsynligheten for å få en falskt positiv verdi 6,5% ($1 - 0,9973^{25}$) i et kontroll-diagram med 25 punkter. Hvis man for eksempel bruker en kontrollgrense på 1,96 SD (tilsvarende $p < 0,05$ for hvert punkt) er det 5 % sjanse for at et datapunkt ligger utenfor kontrollgrensene på grunn av tilfeldig variasjon, og med 25 datapunkter vil sjansen for en falsk positiv verdi bli svært høy.

Dette viser at det kan være fornuftig å bruke 3 SD som kontrollgrenser (på samme måte som man har valgt p-verdi på 0,05 som grense for signifikans i vanlig statistikk). Da er sjansen ganske liten for at man tolker noe som egentlig er en normal variasjon som en spesiell variasjon. Men man kunne ha valgt andre og lavere grenser, men det ville øke risikoen for at man tolker vanlig, tilfeldig variasjon som spesiell variasjon og setter i gang tiltak som kan vise seg å være helt unødvendige og noen ganger uheldige.

Det er vanlig i et kontroll-diagram også å tegne inn grensene for 1 og 2 standardavvik fra gjennomsnittet (altså ikke bare 3 SD som er kontrollgrensene).

5.2.2 Hvordan oppdage spesiell variasjon

Vanligvis vil følgende 4 tester på spesiell variasjon være adekvate for de fleste prosesser i helsevesenet og være en middelvei mellom det å finne falsk spesiell variasjon (type-I-feil) og det å miste spesiell variasjon (type-II-feil).

Test 1 Spesiell variasjon er tilstede når et enkelt punkt faller på utsiden av en kontrollgrense.

Dette var den eneste testen Shewart benyttet. Han brukte kontrollgrense som lå 3 standardavvik (sigmas) fra senterlinjen. Dette var en pragmatisk beslutning, som senere eksperter har fastholdt. Hvis dataene er normalfordelt er sannsynligheten liten (0,27 %) for at et punkt ligger utenfor kontrollgrensen (3 SD) ved tilfeldighet når prosessen ikke er endret.

Test 2: Spesiell variasjon er tilstede når 2 av 3 etterfølgende verdier er på samme side av senterlinjen og mer enn 2 standardavvik fra senterlinjen

² Standardavvik eller standarddeviasjon (SD) er et matematisk mål for hvor langt enkeltdataene ligger fra gjennomsnittet. Det er altså et mål for spredningen.

Denne testen krever at man ikke bare tegner opp kontrollgrensen (3SD) i diagrammet, men også 2 SD.

Test 3: Spesiell variasjon er til stede når 8 eller flere etterfølgende punkter faller på samme side av senterlinjen.

Ofte brukes denne, fordi den er så enkel, som første test etter test 1. Den vil avsløre et mindre, men konsistent skifte i prosessen som både test 1 og 2 ofte overser.

Test 4: Spesiell variasjon er tilstede ved en trend av 6 eller flere verdier i en rekke som enten øker eller synker.

De fleste eksperter krever minst 6 verdier ved færre enn 20 målepunkter eller minst 7 verdier hvis det er 20 eller flere målepunkter.

5.2.3 Finnes det situasjoner der testene for spesiell variasjon bør modifiseres?

Testene ble utviklet for industrielle formål hvor hovedhensikten var å unngå "falsk alarm" som kunne medføre en stengning av hele prosessen. I helsesammenheng kan det hende at man av hensyn til pasienters liv og helse må bruke grensen på 2 SD (i stedet for 3 SD) som "grense for tidlig varsel" eller bruke 6 punkter i stedet for 8 i test 3.

Testene som er beskrevet er imidlertid de som nå vanligvis brukes innen forbedringsarbeid i helsetjenesten.

Det finnes i tillegg andre tester for spesiell variasjon som brukes i mindre grad, for eksempel:

- 4 av 5 etterfølgende punkter mer enn 1 standardavvik fra senterlinjen.
- Fjorten suksessive punkter som går opp og ned annen hver gang i et sagbladmønster.

5.2.4 Hvor mange datapunkter kreves for et kontroll-diagram?

Et kontroll-diagram bør ha omtrent 20 til 25 datapunkter for å utføre de fire testene som er beskrevet her. Hvis man har færre enn 20 datapunkter, er det større fare for ikke å få med spesielle variasjoner (type-II-feil). Men hvis man finner en spesiell variasjon med færre enn 20 punkter, bør man undersøke den. Med mer enn 30 datapunkter er det økt fare for å finne spesiell variasjon pga tilfeldigheter (type-I-feil). For eksempel vil 3 av 1000 målinger normalt ligge utenfor kontrollgrensene ved en stabil prosess uten spesiell variasjon.

5.3 Å velge passende kontroll-diagram

5.3.1 Identifisere type data

Valg av passende kontroll-diagram starter med å identifisere type data. Det er i utgangspunktet to typer data:

1. Måle-data (også kalt "continuous" eller "variable" data)
2. Telle-data (også kalt "discrete" eller "attribute" data)

Måle-data kan inneha forskjellige verdier på en kontinuerlig skala (eks: ventetid i minutter, varighet av noe, kroppsvekt i gram/kg, lengde av opphold i dager etc).

Telle-data er telt, og ikke målt på en skala. Det kan være av to typer:

A. Avvikende enheter (Nonconforming units)

Avvikende enhet (defekt enhet) betyr at en enhet enten oppfyller eller ikke oppfyller et kriterium.

Disse dataene er dikotome eller binomiale – enten/eller, ja/nei-type data. Det kalles også ”defective units”. Telleren (for eksempel antall pas. som falt) er en del av nevneren (antall utskrevne pas.).

B. Avvik, feil (Nonconformities)

En annen tilnærming til telldata er å telle ikke bare avvikende enheter, men alle avvik (feil/defekter). For eksempel: i stedet for å telle om en pasient har falt eller ikke i løpet av et opphold (en avvikende enhet), kan man telle hvor mange ganger en pasient har falt (avvik). Slik kan altså telleren (antall fall) bli større enn nevneren (for eksempel antall pasientdager).

5.3.2 Om å bestemme hvilket diagram som passer best

Man bør konsultere beslutnings-treet for kontroll-diagrammer (figur 8) når man skal bestemme hvilket diagram som passer best. Dette treet viser ikke alle typer kontroll-diagrammer som finnes (som for eksempel NP-diagram, X-bar- og R-diagram). Diagrammet er laget enkelt, og de 5 diagram-typerne som er tatt med, vil være tilstrekkelige for de fleste situasjoner i helsevesenet.

5.3.3 Om subgrupper

Subgruppe er observasjoner produsert av prosessen og som skal presenteres som ett datapunkt i et kontroll-diagram. En subgruppe kan være alle observasjoner over en gitt tidsperiode (for eksempel en time, en dag eller en uke), det kan være et utvalg av de samme observasjonene (for eksempel hver 10. observasjon) eller det kan være en enkelt observasjon (for eksempel antall pasienter en dag).

Valg av subgruppe vil avgjøre hva slags kontroll-diagram man bør velge.

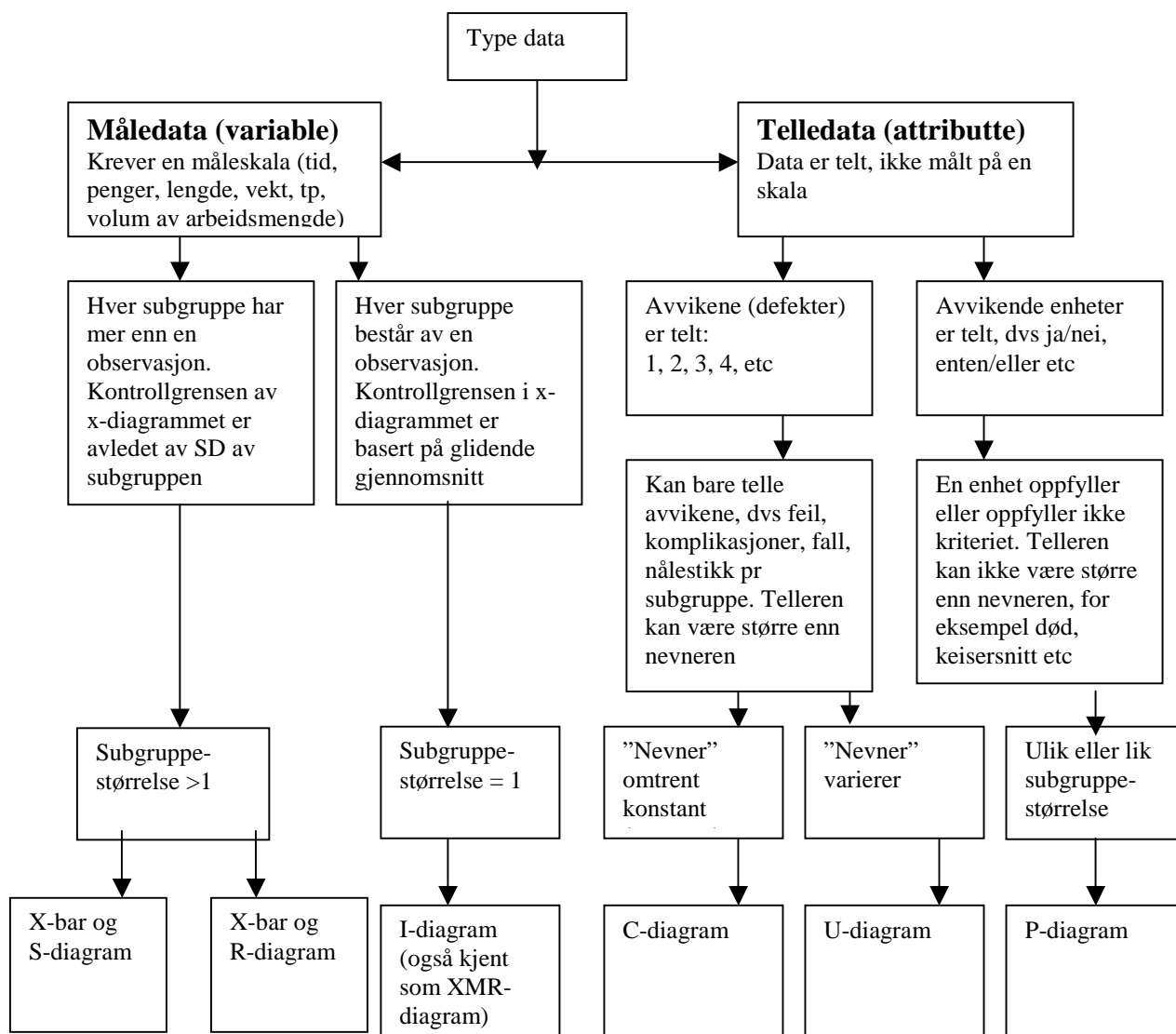
5.4 Ulike kontroll-diagrammer

5.4.1 P-diagram.

P-diagrammet (P står for prosent eller proporsjon) er det diagrammet som er lettest å forstå og kanskje det kontroll-diagrammet som er mest brukt. Det er diagrammet å foretrekke når dataene er telldata av avvikende enheter, for eksempel dødsfall eller keisersnitt. Subgruppen vil bli enheten langs horisontalaksen, som kan være måneder, mens prosent mortalitet eller keisersnitt vil bli plottet på den vertikale aksene. Nevneren vil oftest være ulik fra måned til

måned, men ved å plote prosent i stedet for absolutt tall av avvikende enheter, vil man justere denne differansen.

Eksempel på p-diagram på side 30.



Figur 8. Beslutningstre for kontroll-diagrammer (etter R. Carey):

5.4.2 U-diagram

Der P-diagrammet teller *avvikende enheter* (en enhet kan enten være avvikende eller ikke avvikende), brukes U- og C- diagrammet der hvor hver enhet kan fremvise ett eller flere avvik. U-diagrammet (u står for "**unequal area of opportunity**") er passende diagram når man teller avvik hvor det er varierende utvalgsstørrelse. For eksempel hvis man teller antall fall på en avdeling over en periode, er det oftest ulikt antall pasientdager fra periode til periode. For å gjøre opp for denne ulikheten i mulighetene for fall, deler man antall fall med det totale antall pasientdager i hver periode slik at man får en ratio av fall i forhold til pasientdager. Dette er en mer kraftfull måte å avsløre endring i pasientsikkerhets-prosessen enn med p-diagrammet som viser prosent av pasienter som har falt i hver periode.

5.4.3 C-diagram

Et c-diagram (c står for "constant area of opportunity") er et alternativ til u-diagrammet for tellinger av avvik når det er lik (eller tilnærmet lik) utvalgsstørrelse. Når nevneren er omtrent den samme, dvs ikke varierer mer enn om lag 20 %, kan man plote det aktuelle tallet for hendelser (for eksempel antall fall).

I praksis brukes ofte u-diagram og c-diagram sammen for å vise ulike perspektiver på prosessen.

Eksempler på c-diagram på side 28 og 29.

5.4.4 I-diagram (= XMR-diagram)

Står for "individuelle verdier og glidende gjennomsnitt". Kalles også XMR-diagram (X står for individuell verdi, MR for "moving range"). Det er altså egentlig et dobbeltdiagram, ett med individuelle verdier (X-diagram) og ett diagram med glidende variasjonsbredde (moving-range – MR-diagram). Carey mener imidlertid at det vanligvis ikke er nødvendig å presentere MR-diagrammet. Når hver subgruppe (hvert datapunkt) består av en enkel observasjon, er I-diagrammet det diagrammet som passer best. Selv om det er beregnet på bruk med måledata, kan det også ofte brukes med telldata. Wheeler har hevdet at i-diagram kan brukes som standard også ved telldata. Carey går ikke så langt og skriver: "A beginner in the use of SPC might follow this recommendation." Han viser til at telle-diagram noen ganger kan påvise ekte spesiell variasjon som i-diagram ikke påviser. Jfr diagrammene 2 og 4 i eksempelsamlingen i dette kompendiet.

Eksempler på i-diagram finnes det flere av i kapittel 7.

5.4.5 X-bar og R-diagram

Hvis målingene er gjort slik at subgruppen består av flere enn 1 observasjon og man ønsker å få framstilt hvor stor spredning det er på verdiene i subgruppen, kan man bruke X-bar og R-diagrammet, hvor X-bar står for gjennomsnitt og R står for "range", altså spredning. X-bar-diagrammet framstiller variasjonen mellom subgruppene over tid, mens R-diagrammet viser spredningen.

Carey mener at man kan droppe dette diagrammet og i stedet alltid bruke X-bar og S-diagram når subgruppen er større enn 1. Vi synes imidlertid det noen ganger kan være nyttig å få fram de ekstreme verdiene, for eksempel ved målinger av ventetider hvor vi da kan se hvor lang maksimal ventetid har vært i subgruppen. Vi har derfor valgt å ta med dette diagrammet.

5.4.6 X-bar og S-diagram

Hvis målingene er gjort slik at hver subgruppe består av flere enn en observasjon (tidligere ble det sagt flere enn 10 observasjoner), kan man bruke X-bar og S-diagrammet. X-bar står for gjennomsnitt og S står for standardavvik. Et eksempel på anvendelse er å måle om lengden av opphold ved et sykehus har endret seg over tid. Her registreres alle (eller et utvalg av) alle målinger i hver subgruppe. S-diagrammet framstiller variasjonen i hver subgruppe, mens X-

bar-diagrammet (som betyr gjennomsnittsdigrammet) framstiller variasjonen mellom subgruppene over tid.

Eksempler på X-bar og S-diagrammer på side 32 og 33.

5.5 Hvilket kontroll-diagram er best?

Det passende diagram for en situasjon avhenger av flere faktorer. Generelt kan man si :

- kontroll-diagrammer for målinger er mer kraftfulle enn kontroll-diagrammer for telldata
- X-bar og S-diagram er mer kraftfullt enn I-diagram
- C- diagram og U-diagram er mer kraftfulle enn P-diagram.

Derfor bør man forsøke å finne det beste diagrammet, ikke bare et korrekt diagram.

5.6 Subgruppestørrelse

5.6.1 Subgruppestørrelse for p-diagrammer

En foreslått retningslinje for dette er at subgruppe-størrelsen skal være så stor at man finner minst en hendelse i hver subgruppe. Det skal altså ikke være noen subgruppe med 0 % hendelser.

En annen regel som er foreslått er at subgruppene skal være så store at kontroll-diagrammet har en positiv nedre kontrollgrense.

5.6.2 Subgruppestørrelse for c-diagrammer og u-diagrammer

En foreslått regel er at subgruppe-størrelsen for et u-diagram bør være minst 1 delt på gjennomsnittet av avvikene (\bar{u}). For c-diagrammer bør subgruppe-størrelsen være stor nok til at gjennomsnittet av avvikende enheter (\bar{c}) er større enn 1, men helst større enn 4.

6 Appendix:

6.1 Diverse begreper

Run chart	=	Run-diagram, forløpsdiagram
Control chart	=	Kontroll-diagram
Key quality characteristic	=	viktig kvalitetsegenskap (alternativt begrep: nøkkelmål for kvalitet)
Nonconforming units	=	avvikende enheter
Nonconformities	=	avvik, avvikshendelser, defekter
Moving range	=	glidende endring (glidende variasjonsvidde)
Common cause variation	=	Naturlig variasjon, ev. tilfeldig variasjon, vanlig variasjon, variasjon basert på konstante, iboende årsaker (variasjonen er en innebygget egenskap ved prosessen)
Special cause variation	=	Spesiell variasjon, variasjon basert på spesielle årsaker (denne variasjonen har en konkret årsak og kan fjernes ved at årsaken finnes og elimineres)
Øvre/nedre kontrollgrense	=	Viser yttergrensene for naturlig variasjon (3 SD)
Prosessens kapabilitet	=	Spredningen (mellom øvre og nedre kontrollgrense)
Run	=	løp (sekvens av påhverandrefølgende verdier som alle ligger på samme side av prosessens medianverdi)
Trend (eng.)	=	trend (norsk)
Senterlinjen, sentrallinjen	=	I kontroll-diagrammer: gjennomsnittet som viser prosessens gjennomsnittsyttelse (Aune: middelerdi for prosessen) I run-diagram: medianverdien
Outliers	=	Sporadiske avvik (uteliggere) (Aune: ensomme ulver!)
Type 1-feil	=	Normal variasjon som tolkes som en spesiell variasjon
Type 2-feil	=	Hvis man mister en spesiell variasjon

6.2 Litteratur

6.2.1 Bøker/hefter/tidsskrifter/tidsskriftsartikler

Carey RG. "Improving healthcare with control charts. Basic and advanced SPC methods and case studies". ASQ Quality Press, 2003

Carey RG. and Lloyd RC. "Measuring Quality Improvement in Healthcare: A Guide to Statistical Process Control Applications." ASQ Quality Press, 2001

Hart MK and Hart RF: "Statistical Process Control for Health Care". Duxbury Applied Series, 2002

Kvalitetsforbedring i psykisk helsevern: Prosessforbedring i klinisk virksomhet. Oslo, Helsetilsynet, 2001. 63s. IK-2756.

InFORUM – Tidsskrift for Norsk Forum for Kvalitet i Helsetjenesten (NFKH)

Wheeler DJ: "Understanding Variation: The Key to Managing Chaos", 1993.

Benneyan JC, Lloyd RC, Plsek PE: "Statistical process control as a tool for research and healthcare improvement", Qual Saf Health Care 2003; 12:458-464

Wheeler DJ: "Advanced Topics in Statistical Process Control" SPC Press, 1995

6.2.2 Dataprogrammer

CHARTrunner 2000 <http://www.chartrunner.com/>

Minitab <http://www.minitab.com/>

QI Analyst <http://www.wonderware.com/products/analytical/qianalyst>

QI Macros <http://www.qimacros.com/Macros>

Quality America <http://www.qualityamerica.com/>

Statit <http://www.statit.com/>

7 Eksempelsamling

7.1 Run-diagram

Ved legevakten i Skien har man gjennom en periode registrert hvem pasientene som kommer til konsultasjon på kveldstid har som fastlege, da det har vist seg at noen legekontorer ”rekrutterer” dobbelt så mange pasienter til legevakta enn andre, også etter at tallene er justert for listelengde. Vi mener at tallene sier noe om hvordan det enkelte legekontor betjener sine pasienter i forhold til behov for akutt legehjelp.

Registreringene går tilbake til legene. Tallene for 27 måneder for et en-legekontor er her satt opp i et rundiagram.

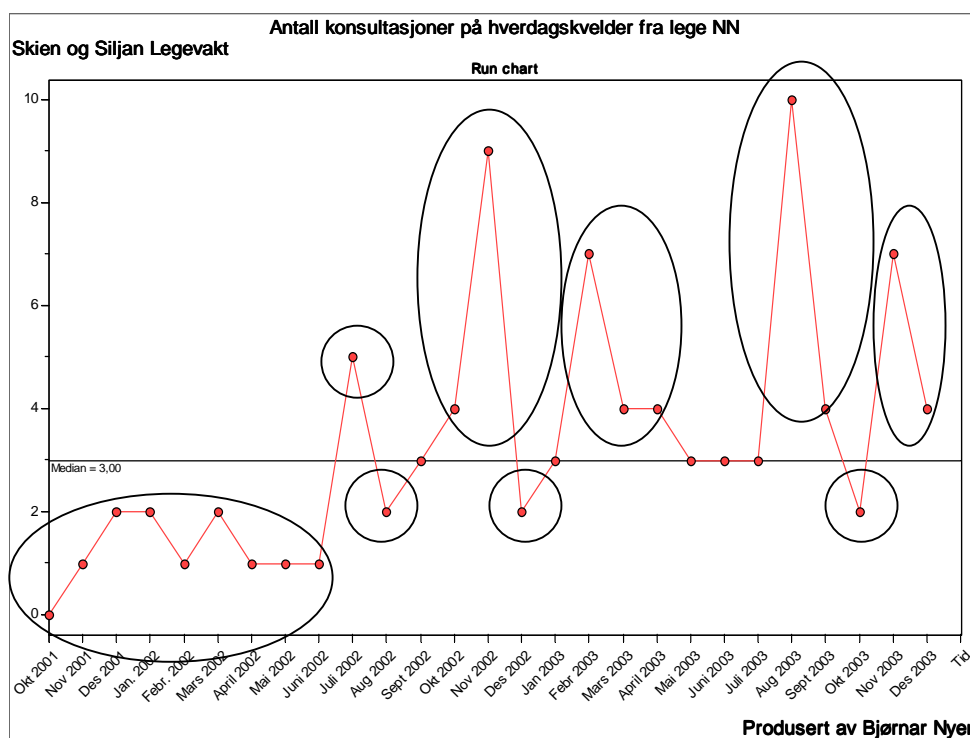


Diagram 1 *Antall kveldskonsultasjon av pas. fra lege NN fra oktober 2001 til juli 2003 (run-diagram)*

Hvor mange runs finner du her? (kfr side 16)

Finner du spesiell variasjon? (kfr side 17)

7.2 I-diagram

De samme tallene som i Figur 1 er lagt inn i et I-diagram (diagram 2).

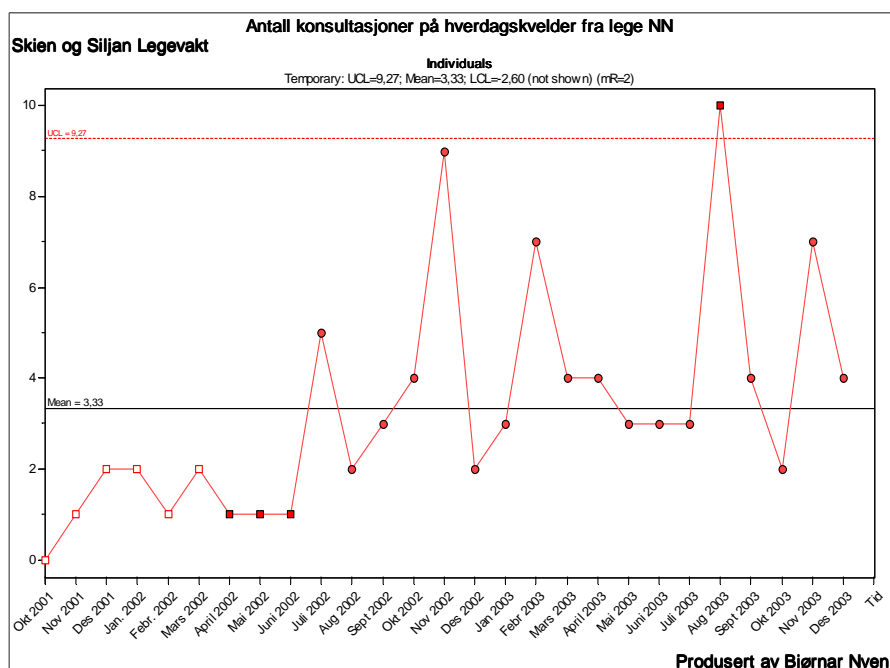


Diagram 2 *Antall konsultasjoner på hverdagskvelder ved legevakt av pasienter fra fastlege NN (I-diagram)*

Finner du spesielle variasjoner her? (kfr side 18 og 19)

Her ser vi et eksempel på at et i-diagram avslører spesiell variasjon som run-diagrammet ikke avslører.

Diagram 3 viser resultatet fra samme registrering for alle legene ved et fastlegekontor.

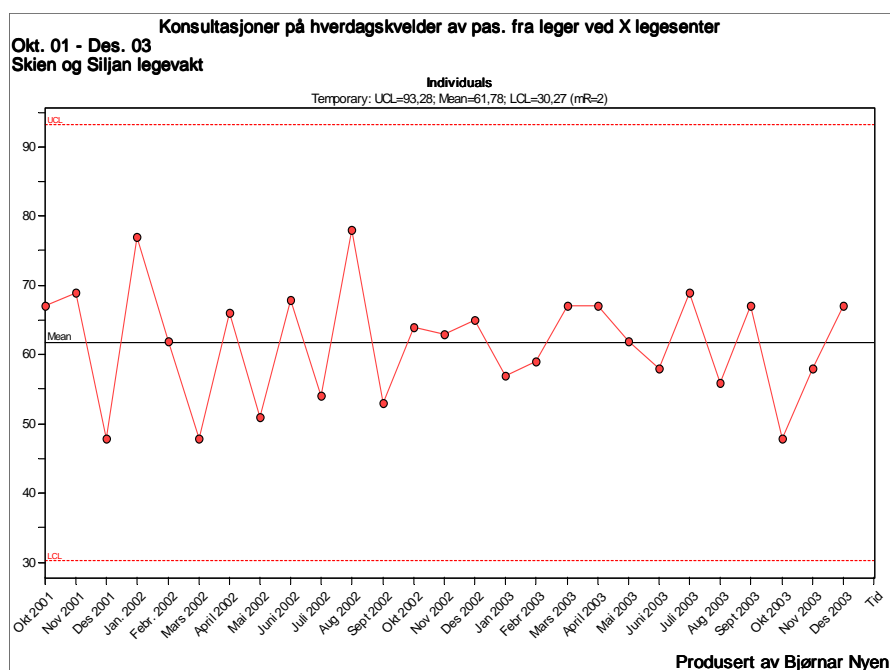


Diagram 3. *Kontrolldiagram – antall kveldskonsultasjon av pas. fra Legesenter X fra okt 2001 til juli 2003 (I-diagram)*

Hvordan vil du bedømme dette diagrammet?

7.3 C-diagram

Tallene fra denne registreringen av antall konsultasjoner på legevakta er telldata. Siden ”nevneren” (legens/legenes pasientliste) er omtrent konstant fra måned til måned, kan man velge c-diagram (c- står for ”constant area of opportunity”), diagram 4.

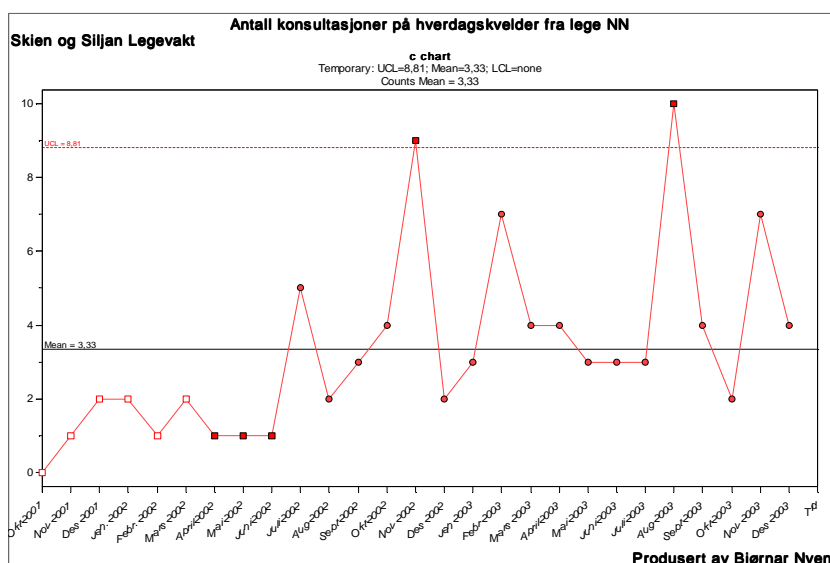


Diagram 4 *Antall konsultasjoner på hverdagskvelder ved legevakt av pasienter fra fastlege NN (C-diagram)*

Her dukker enda en spesiell variasjon opp. Altså kan valg av type diagram avgjøre hva vi finner av variasjon.

7.4 Ubesvarte telefoner – I-diagram og P-diagram

Ved et legekantor i Skien har man i 3 uker registrert alle innkommende telefoner på ventetid og på om det blir svart. Resultatene for de tre ukene er lagt sammen, slik at en times registrering i diagrammet er uttrykk for situasjonen i den timen i tre fortløpende uker.

Diagram 5 viser antall besvarte telefoner (c-diagram). Diagram 6 viser antall ubesvarte telefoner (c-diagram).

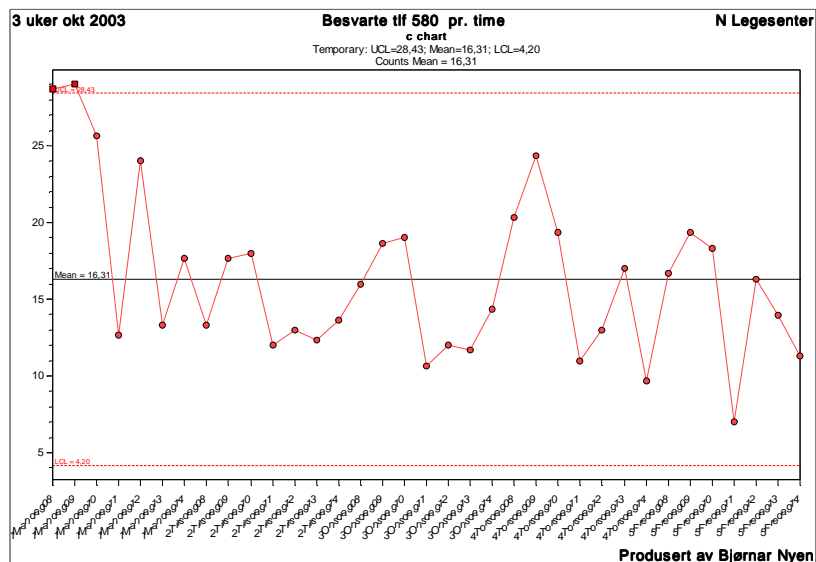


Diagram 5 *Antall besvarte telefoner (gjennomsnittlig per time) ved N legesenter (I-diagram)*

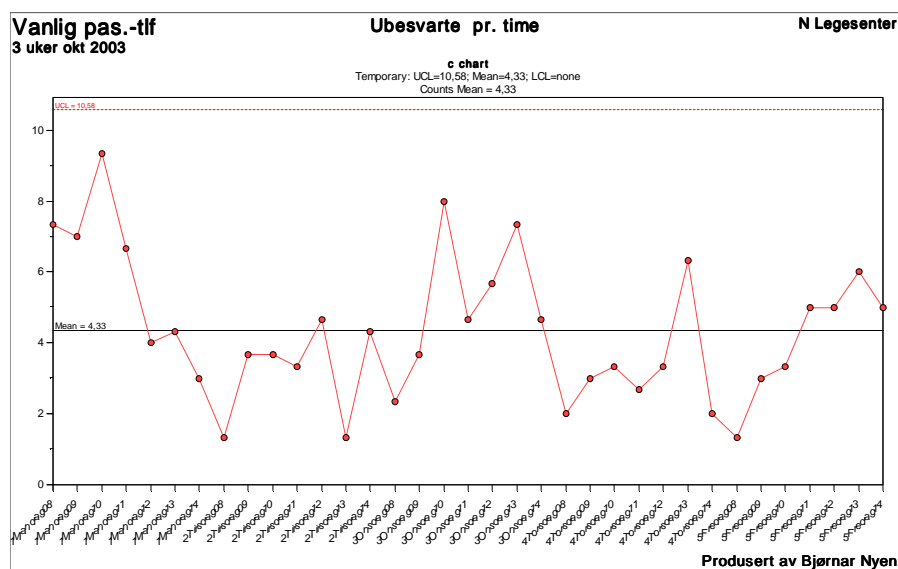


Diagram 6 *Antall ubesvarte telefoner (gjennomsnittlig per time) ved N legesenter (C-diagram)*

Hvis dataene fra diagram 6 legges inn i et i-diagram, da får vi fram en spesiell variasjon mandag formiddag (diagram 7).

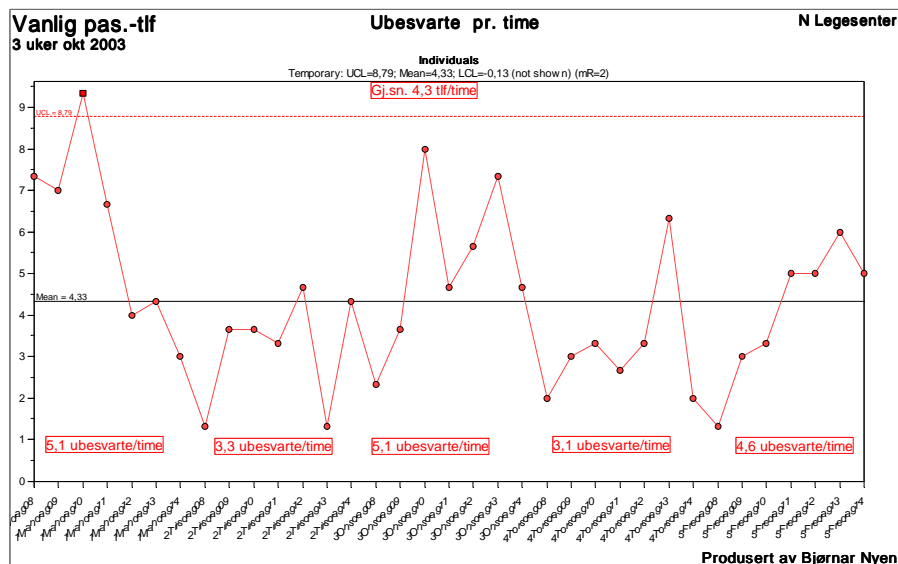


Diagram 7 Antall ubesvarte telefoner (gjennomsnittlig per time) ved N legesenter (I-diagram)

Diagram 8 (p-diagram) uttrykker forholdet mellom antall ubesvarte og totalt antall innringninger (besvarte og ubesvarte) time for time.

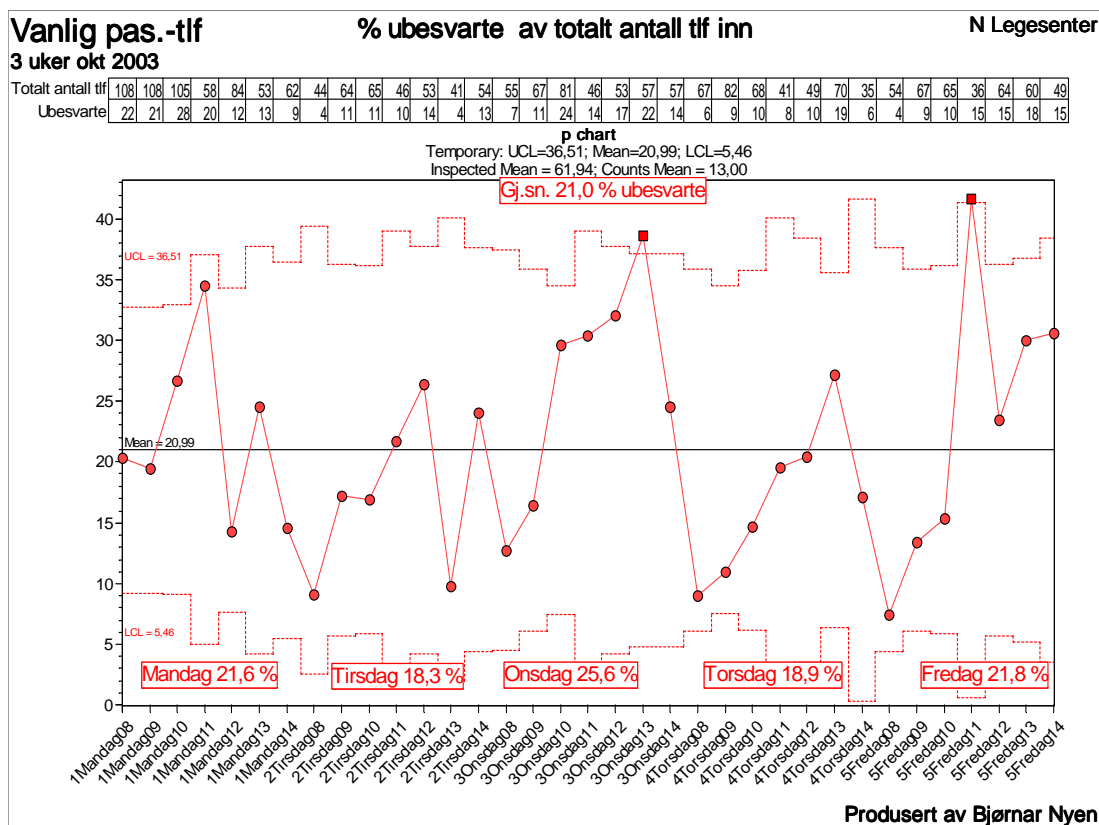


Diagram 8 Andel ubesvarte telefoner ved Legesenter N målt over 3 uker i okt 2003 (P-diagram)

Er prosessen stabil?

Har du synspunkter på nivået (gjennomsnittlig 21 % ubesvarte telefoner)?

7.5 Ekstra mye å gjøre på legevakta – I-diagram

Ofta opplever vi at det er alt for mye å gjøre – pasientene synes å strømme på. Det er sjelden vi trenger statistikk for å oppdage dette! Men noen ganger er det bra å påvise endringer også. I dette tilfelle kunne statistikken ha blitt brukt til å forberede seg på en travel periode. Dessuten er den i etterkant blitt brukt til å rettferdiggjøre bruk av ekstra personell.

Vi har registrert alle kontakter som fører til et notat i journalsystemet ved legevakta. Her er data for alle hverdager i perioden 01.10.03 – 04.01.04

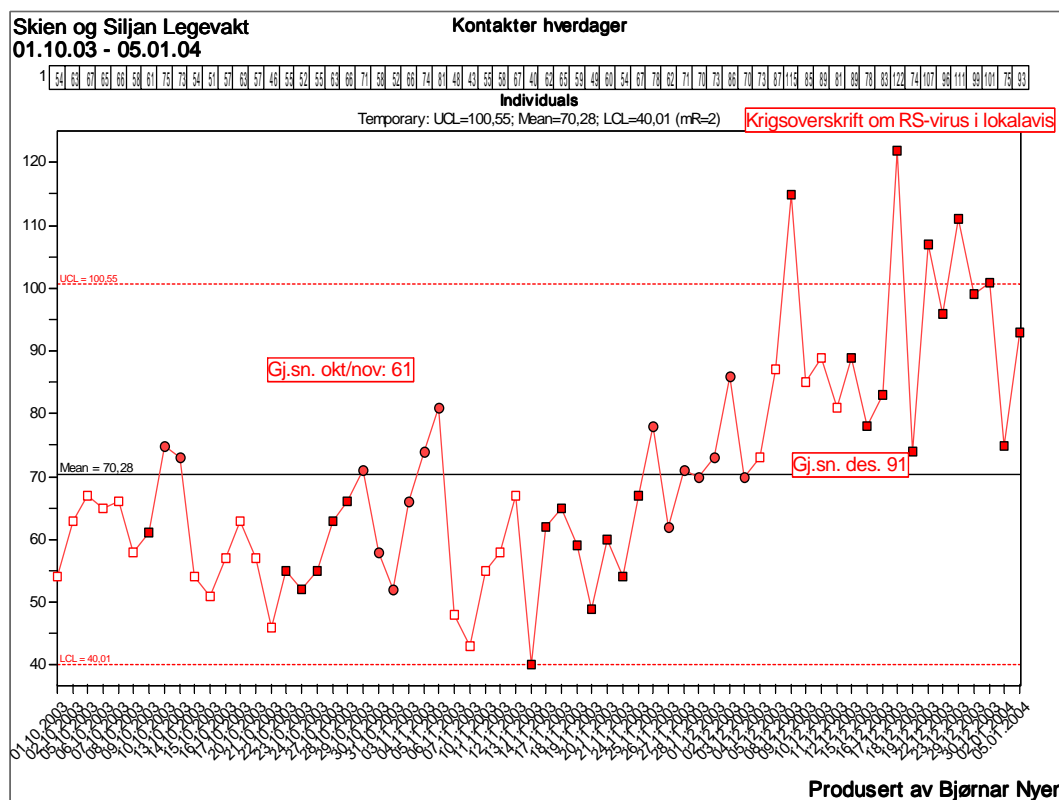


Diagram 9 Kontakter på legevakter hverdager (I-diagram)

Det er interessant å se toppen 17. desember. Den dagen hadde en lokalavis (TA) følgende oppslag på 1. sida: "BARN rammet av farlig virus", og VG hadde dette på 1.sida: "FEM UNGE kan være influensa-DØDE". Vi tror det var spesielt mange som søkte til legevakta den dagen på grunn av de oppslagene, men søkningen var over øvre kontrollgrense 3 andre dager i desember uten at det var liknende avisoverskrifter, så det kan ha være at toppen var helt tilfeldig.

Diagram 10 er samme diagram hvor gjennomsnittet for oktober og november som "baseline" er tegnet inn. Dersom disse dataene hadde blitt fulgt daglig, ville man etter å ha lagt inn datapunktet for 2.12. se at det hadde skjedd mer enn tilfeldig variasjon i antall kontakter (7 datapunkter over gjennomsnittsverdien = spesiell variasjon). Om det var blitt gjort kunne de kanskje ha forberedt seg på den spesielt store pågangen som de hadde ved legevakta utover i desember.

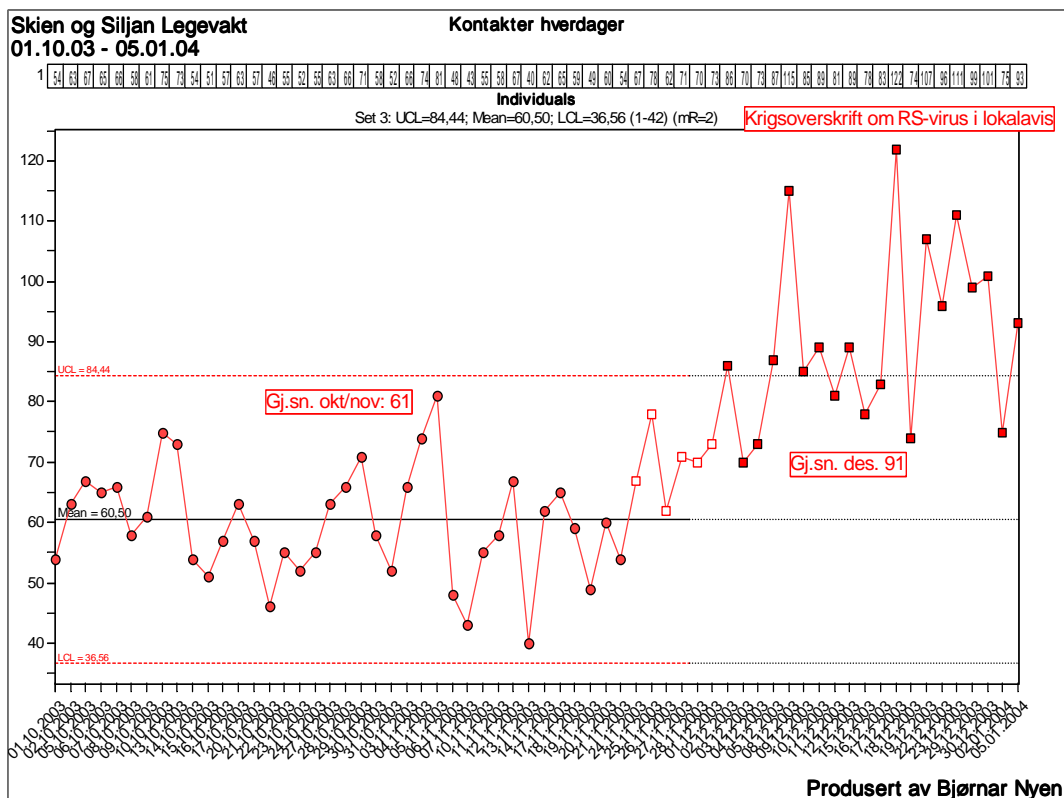


Diagram 10 Kontakter på legevakter hverdager (I-diagram)

7.6 Ventetid på legevakt X-bar S-diagram

Ved legevakten i en by på Østlandet har man sett på ventetiden fra en pasient henvender seg i ekspedisjonen til konsultasjonen hos lege starter. Her er resultatene for 30 dager satt opp i et X-bar og S-diagram.

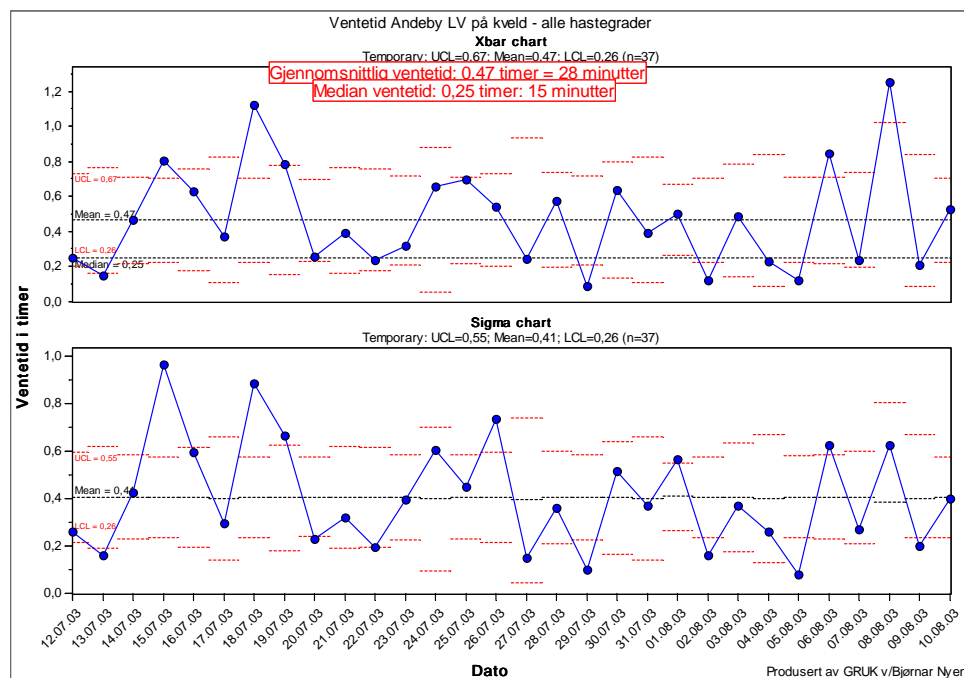


Diagram 11 Ventetid ved "Andeby legevakt" alle hastegrader 120703 – 100803 (X-bar S-diagram)

Hvilke kommentarer har du til dette?

Stabilitet?
Akseptabel ventetid?

7.7 Ventetid på telefon – X-bar R-diagram og X-bar S-diagram

Det er gjort telefontrafikk-måling ved et fastlegekontor hvor man også har sett på ventetid før innringer får svar. Dataene er lagt inn i et X-bar S-diagram (diagram 12).

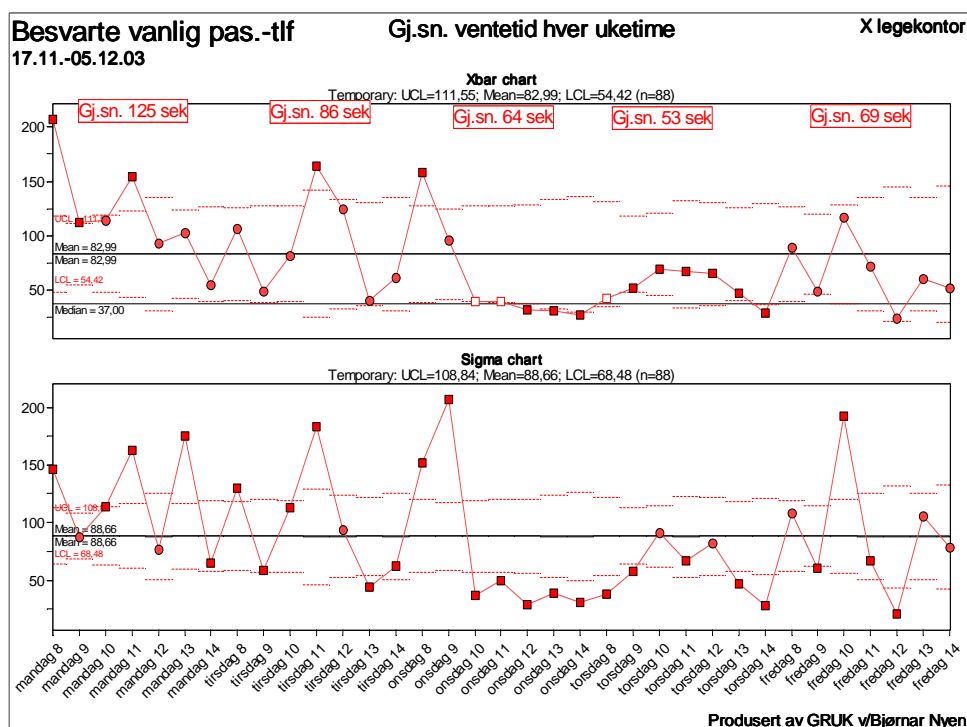


Diagram 12 Ventetid på telefon – besvarte telefoner time for time gjennom arbeidsuka (gjennomsnitt for 3 uker) (X-bar S-diagram)

De samme dataene er lagt inn i et X-bar R-diagram (diagram 13). I det diagrammet kan vi se at innringere har ventet bortimot 1200 sekunder (20 minutter) før de fikk svar.

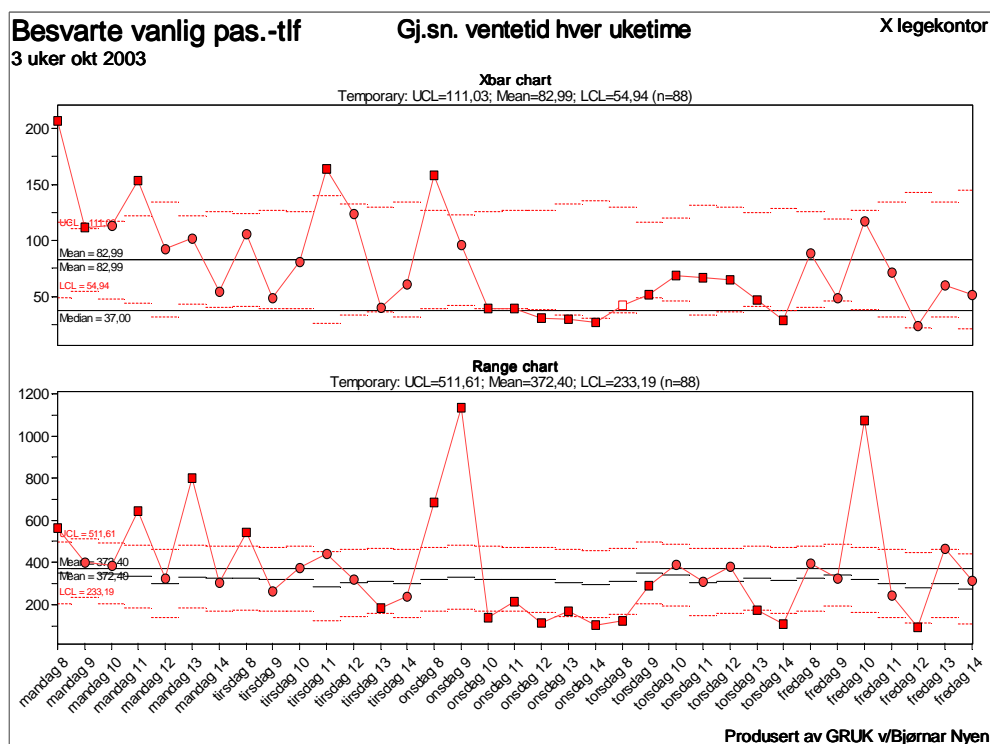


Diagram 13 Ventetid på telefon – besvarte telefoner time for time gjennom arbeidsuka (gjennomsnitt for 3 uker) (X-bar R-diagram)

7.8 INR og Marevan-dosering

Her følger en historie om Marevan-dosering som kan illustrere at det kan være nyttig å forstå variasjon. Eksemplet er tatt fra Careys bok “Improving healthcare with control charts. Basic and advanced SPC methods and case studies”.

En mann på 72 år var marevanisert på grunn av klaffefeil. INR-mål var satt til 2,5 – 3,5. Dosering ble styrt etter INR-målingene ved at om INR var lavere enn 2,5 ble dosen økt, og om INR var høyere enn 3,5 ble dosen redusert.

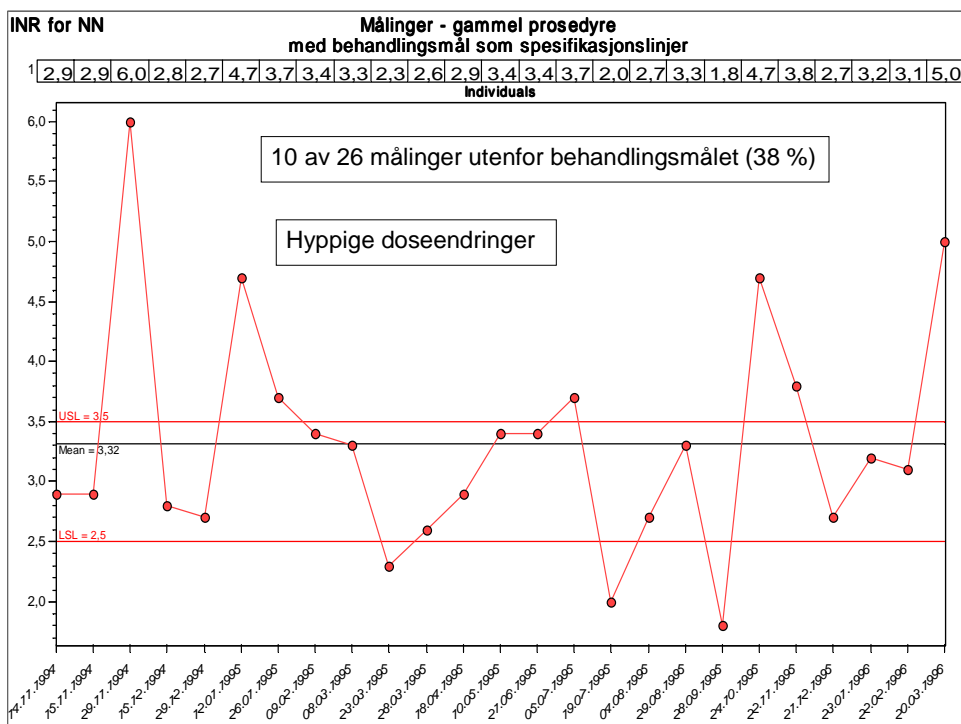


Diagram 14 INR-målinger med gamle retningslinjer. Intervensjonsgrensene (sammenfallende med behandlingsmål) inntegnet (I-diagram)

Så ble behandlingsstrategien endret slik:

- Behandlingsmål på 2,5 – 3,5 beholdes
- Ved verdier utenfor behandlingsmålet, gjøres ny måling (uten endring av dosen)
- Bare hvis verdien er utenfor øvre behandlingsmål + 0,5 (dvs 4,0) reduseres dosen
- Bare hvis verdien er under nedre behandlingsmål + 0,5 (dvs 2,0) økes dose

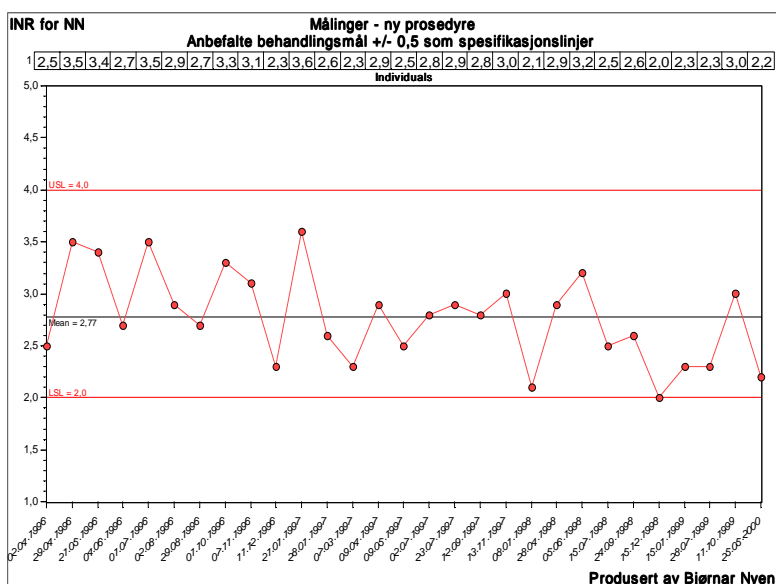


Diagram 15 INR-målinger med nye retningslinjer. Intervensjonsgrensene inntegnet (I-diagram)

7.9 Mer om spesiell variasjon

Når man finner en spesiell variasjon i en prosess bør man forsøke å finne årsaken til denne variasjonen. Om man finner årsaken vil man noen ganger kunne bygge om prosessen for å unngå liknende variasjon senere.

Andre ganger kan det være en spesiell årsak som ikke er ønskelig, men som vi mener vi ikke kan unngå. Her følger et eksempel på det. Ved en legevakt vil pågangen variere betydelig. Det er ikke rimelig av samfunnsøkonomiske grunner å dimensjonere legevakta ressursmessig slik at det aldri blir noe ventetid. Ved stor pågang må personellet prioritere hva de skal ta først, og hva som kan vente. Ved en registrering av ventetid for pasienter som blir innlagt i sykehus for magesmerter, fant man følgende (diagram 18):

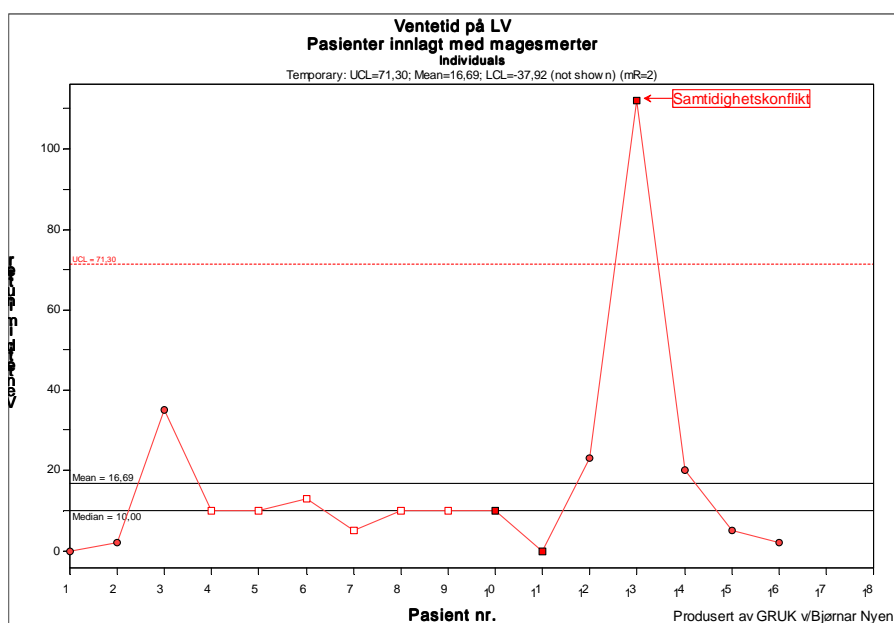


Diagram 18 Ventetid på legevakta for pasienter som senere ble innlagt i sykehus for magesmerter (I-diagram)

Pasient nr. 13 måtte vente i 112 minutter fra ankomst til legevakta til konsultasjon hos lege startet. Grunnen til den betydelig lengre ventetiden for denne pasienten enn andre med magesmerter, var at det kom skadete fra en bilulykke omtrent samtidig, som måtte prioriteres før denne pasienten. Dessuten kom det en annen pas. (pas. nr. 14) med magesmerter, som av sykepleier ble bedømt til å være mer kritisk syk enn pasient nr. 13 og derfor fikk konsultasjon hos lege før pas. nr. 14 selv om hun kom til legevakta etter.

Det er ingen grunn til å ta ut pas. 13 fra datamaterialet når man regner gjennomsnitt og kontrollgrenser, fordi legevakta må regne med å få liknende tilfeller også senere, selv om det ikke er ønskelig.

I diagram 19 er det registrert en pasient med betydelig lengre behandlingstid på legevakta enn de andre – pas. nr. 8 er registrert med en behandlingstid på 160 minutter som ligger godt utenfor kontrollgrensen i dette diagrammet. Grunnen til den lange behandlingstiden, var at personellet ved legevakta ønsket å observere pasienten over noe tid før beslutning om tiltak ble tatt. Det er derfor rimelig å ta denne pasienten ut av datamaterialet når man regner gjennomsnitt og kontrollgrenser for behandlingstiden på legevakta.

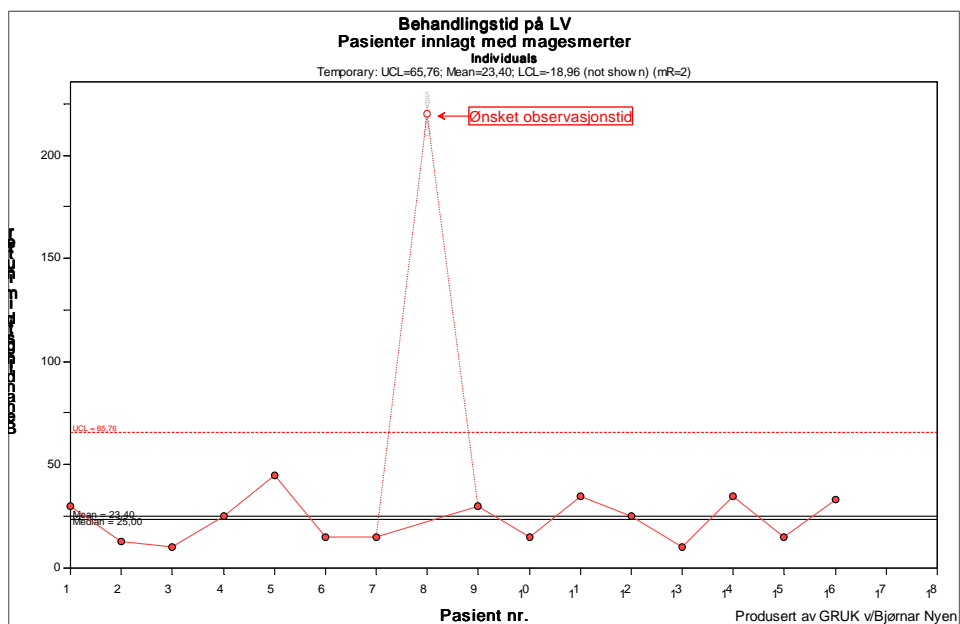


Diagram 19 *Behandlingstid på legevakt for pasienter som ble innlagt i sykehus for magesmerter (I-diagram)*