

Standardisering av resultater

Offentliggjøring av resultater fra medisinske kvalitetsregistre

www.kvalitetsregistre.no/standardisering-av-resultater-til-offentliggjøring

For en enhetlig fremstilling og et bedre totalinntrykk baserer offentliggjøringen av resultater fra de medisinske kvalitetsregistrene seg på en standardisering av figurer og tabeller. Dette dokumentet drøfter informasjonsformidling generelt og gir en detaljert beskrivelse av hvordan resultatene fra kvalitetsregistrene skal presenteres ifm. offentliggjøringen. De som finner det generelle stoffet trivielt kan lese oppsummeringen i **Avsnitt 6** og benytte det som utgangspunkt for selektiv lesning av resten av dokumentet, primært **Avsnitt 5**.

Det nasjonale servicemiljøet, hvor hver region har ansvar for tilhørende registre, vil i den grad kapasiteten strekker til, kunne tilby hjelp med å lage figurer. Registerne som kan gjennomføre standardiseringen selv oppfordres sterkt til dette for å lette belastningen på de regionale enhetene. Eksemplene i dette dokumentet er laget med statistikkverktøyet R, men registerne må gjerne bruke andre digitale verktøy så lenge form, farge og filformater er i tråd med retningslinjene.

Innhold

1 Introduksjon	1
2 Filformater	1
3 Farger	1
4 Visualisering	1
4.1 Grunnprinsipper	1
4.2 Valg av figurer og tabeller	2
5 Spesifikke maler	2
5.1 Mal for søylediagram	2
5.2 Mal for linjediagram	3
5.3 Mal for andre figurer	4
5.4 Mal for tabeller	4
6 Oppsummering	4
Noter	4
Appendiks	4
Tabeller	5
R-kode	5

1. Introduksjon

Hensikten med en datavisualisering kan være analyse (søke mønstre) eller presentasjon (formidle funn). Formålet med offentliggjøringen av resultater fra de medisinske kvalitetsregistrene er primært sistnevnte. Man må derfor ta hensyn til publikums evner og forventninger; informasjonen må tilpasses mottakernes behov og kunnskapsnivå.

Utfordringen med offentliggjøringen er at målgruppen er sammensatt; den består ikke bare av medisinske fagmiljøer, men også pasienter og pårørende. Et akademisk/vitenskapelig nivå i resultatformidlingen er derfor ikke hensiktsmessig. Medisinske begreper og faguttrykk bør også unngås der mer «folkelige» alternativer eksisterer.

Målet er å kommunisere data og funn fra kvalitetsregistrene på en klar, effektiv og presis måte. Men dette er avveiningssak, høy presisjon kan f.eks. gå på bekostning av klarhet, og man må trolig innfinne seg med et kompromiss hvor man unngår utstrakt bruk av p-verdier og beslektede begreper. Mottakernes tanker skal ledes til *innholdet fremfor metodikken*.

2. Filformater

Alle figurer skal leveres som egne filer. Etersom figurene skal presenteres digital og være egnet for utskrift holder det ikke å *kun* lime figurene i et samlet worddokument. Grafikkfiler bør være vektorbasert¹, og primært ønskes filformatet PDF, alternativt bildeformatet PNG med **høy oppløsning**, gjerne 300 PPI². Tabeller kan levers i Worddokumenter, regneark (XLS) eller ren tekst (CSV).

3. Farger

Primærfargene i paletten som skal benyttes er seks **blånyanser**, hvor man for hver figur benytter de nyanser som er egnet mtp. formålet (blånyanse nr. 1, 3 og 5 i **Tabell 2** vil f.eks. skille seg bedre fra hverandre enn nr. 2, 3 og 4). Den mørkeste blåfargen kan av hensyn til fargeharmoni, anvendes som et alternativ til sort farge på fonter og lignende (som overskriftene i dette dokumentet). Støttefargene som supplerer blåfargene er fire **grånyanser**. Det er også definert en kontrastfarge som kan benyttes hvis elementer skal utheves spesielt (se **Figur 7**).

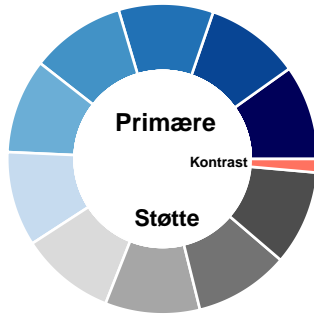
Husk at det er mulig å bruke andre effekter enn ulike farger for å illustrere forskjeller, f.eks. stiplede linjer og ulike punkttyper. Det kan likevel oppstå tilfeller hvor graderinger av en ekstra farge er nødvendig. Da kan turkis, eller annen harmoniserende/passende farge benyttes. De figuransvarlige har altså en viss kunstnerisk frihet – gjør de valg som fungerer best mtp. figurenes leselighet uten at de blir for fargeskrikende.

Paletten er illustrert i **Figur 1** på neste side, og fargekodene³ er gitt i **Tabell 2** på side 5 samt implementert i den vedlagte **R-koden**. Fargene kan også interpoleres ved behov for flere nyanser.

4. Visualisering

4.1 Grunnprinsipper

Kunsten med god informasjonsformidling er å fremheve mønstre i store datamengder ved å presentere mye informasjon på liten plass. Ideelt sett skal visualiseringen gi et godt overblikk samtidig som den muliggjør avlesning av finere strukturer og



Figur 1. Offentliggjørings fargepalett.

detaljer. Man snakker gjerne om å *optimere informasjonsflyten* eller å minimere mottakerens anstrengelser ved å finne relevant informasjon. I beste fall er en grafisk fremstilling selvforklarende, som et minimum er den fullt lesbar med en supplerende figurtekst. Verre er det om mottakeren må lese gjennom lengre tekstavsnitt for å forstå figuren, og verst er det om figuren er uforståelig for målgruppen.

Historisk har man ønsket å minimere blekkforbruk per informasjonsenhet, og selv om vinklingen er mindre relevant med dagens digitale verktøy, er det fortsatt et godt prinsipp å unngå flere farger, effekter, former og markeringer enn det strengt nødvendige. Kreativitet, redundans og visuelt støy (såkalt «chart junk⁴») er i denne sammenheng uønsket – man søker klare budskap, systematisk standardisering og en kompresjon av innholdet ned til det essensielle. Men heller ikke mer; alle figurer skal ha en merking av akser og beskrivelse av eventuelle farger, form og markeringer. Overskrifter samt tabell- og figurtekst skal dog **ikke** inkluderes i grafikkfilene.

4.2 Valg av figurer og tabeller

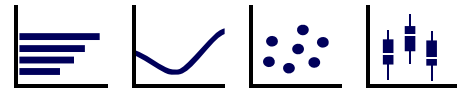
De fire mest brukte figurtypene er søylediagram, linjediagram, spredningsplott og boksplokk, som illustrert av Figur 2. Man får illustrert det aller meste kun ved hjelp av disse. Hvilken figurtype som bør benyttes er avhengig av budskapet som skal formidles og de underliggende datas natur.

Det er hensiktsmessig å skille mellom numeriske og kategoriske variabler, avhengig av om de inntar en numerisk verdi eller en av flere mulige kategorier. Ifm. offentliggjøringen presenteres typisk målinger av en numerisk variabel med en kategorisk tilhørighet (f.eks. overlevelse ved et sykehus). Merk at man kan generere en kategorisk variabel fra en numerisk variabel ved å dele utfallsrommet inn i intervaller.

Søylediagrammet er godt egnet for avlesning av numeriske verdier inndelt i båser/kategorier. Som regel kan nøkkeltall, som referanse- og ekstremalverdier, raskt identifiseres og settes i sammenheng. For sammenligninger av numeriske verdier på tvers av kategorier (som f.eks. sykehus) bør et sortert søylediagram benyttes. I visse tilfeller kan søylediagrammet også anvendes for å illustrere andeler, avvik (f.eks. positivt og negativt avvik fra en verdi/linje) eller fordelinger. Offentliggjøringsmalen for søylediagrammet er gitt i Avsnitt 5.1.

Helseforetak	Region	Verdi
Et HF	Midt	7.5
Et annet HF	Nord	7.1
Enda ett HF	Midt	6.8

Tabell 1. Et enkelt tabelloppsett.



Figur 2. Søyle- og linjediagram, sprednings- og boksplokk.

Linjediagrammet er godt egnet for å illustrere endringer fra punkt til punkt, og benyttes gjerne ifm. tidsserier for å fremheve trender og fluktasjoner. Ifm. offentliggjøringen skal linjediagram generelt anvendes for utvikling over tid. Der som ulike kategoriers/grupper utvikling skal sammenlignes benyttes flere linjer i samme figur. I enkelte spesialtilfeller kan linjediagrammet også anvendes for illustrasjon av avvik eller fordelinger. Mal for linjediagrammet er gitt i Avsnitt 5.2.

Spredningsplottet benyttes primært i søk etter mønstre og assosiasjoner mellom variabler, og er nok derfor mindre aktuelt ifm. offentliggjøringen, men mal for spredningsplottet omtales indirekte i Avsnitt 5.3.

Boksplokket anvendes ved sammenligning av to eller flere datasett mtp. sentralitet, spredning og skjevhet, og er trolig også lite aktuelt ifm. offentliggjøringen. Mal for spredningsplottet omtales indirekte i Avsnitt 5.3.

Andre figurer enn de fire overnevnte kan være hensiktsmessig i visse sammenhenger. I Avsnitt 5.3 gis en kort veiledning for hvordan man roter dypere i den statistiske verktøykassen for å finne den rette figurtypen.

Tabeller er i visse tilfeller å foretrekke. En figur er vanligvis mest effektivt/intuitiv, men hvis tallene som skal presenteres er få, må angis svært presist, eller kun er ment for enkeltvis oppslag, kan en tabell være best egnet. Man søker da et minimalistisk oppsett som eksemplifisert i Tabell 1 og beskrevet i avsnitt Avsnitt 5.4.

5. Spesifikke maler

Essensen av Avsnitt 4.2 er at søylediagram anvendes for verdi-sammenligninger på tvers av kategorier og linjediagram for verdiers utvikling over tid. Se Tabell 3 på side 5 for flere detaljer. Her følger maler for hvordan figurene skal implementeres.

5.1 Mal for søylediagram

Offentliggjøringsprofilen for **søylediagrammet** er illustrert i Figur 3 og tilhørende R-kode er gitt i appendikset. Koden er ikke hugget i stein, det vil i de fleste tilfeller være nødvendig med tilpasninger og skreddersøm. Figuren er kun et veiledende eksempel, man vil neppe anvende både konfidensintervall og spesifisering av antall (N), eller inkludere merkelappen (nederst i høyre hjørne) dersom det ikke er nødvendig.

Hvis søylene er svært få og deres tilhørende tekst er godt lesbar kan et vertikalt søylediagram benyttes, ellers skal søylene alltid være **horisontale**.

Et søylediagram bør være **sortert**; i fallende rekkefølge fra toppen hvis maksimumsverdier skal fremheves og i stigende/fallende rekkefølge hvis minimumsverdier skal fremheves. Hvis stigende/fallende sortering ikke er fornuftig kan annen (f.eks. geografisk) sortering benyttes. Søylar kan i enkelttilfeller erstattes av punktmarkeringer dersom dette letter avlesningen.

Søylediagrammets numeriske akse skal *alltid* inkludere **nullverdien**. Hvis forskjeller skal fremheves kan aksene kuttes, men da skal søylene erstattes av punkter, som illustrert i Figur 4. Merk hvordan et slik **kuttet søylediagram** fremhever kategoriens individuelle verdier relativt Figur 3.

Andre variasjoner av søylediagrammet velges etter formål. Hvis det er flere variabler per kategori/objekt som skal illustreres, benyttes et **gruppert søylediagram**. Et eksempel er gitt i Figur 5, hvor kategoriene (indikatorne) er grupperte. Dette er spesielt hensiktsmessig dersom variablene er **andeler** og formålet er å sammenligne disse innad i kategoriene. Dersom variabler (søyle 1 og 2) eller andeler skal sammenlignes på tvers av kategoriene gjøres grupperingen etter variabel fremfor kategori. Søylene som skal sammenlignes skal stå så nær hverandre som mulig – altså innad i en gruppering.

Et såkalt **stablet søylediagram**, hvor det er en inndeling av hver søyle, skal *kun* anvendes hvis søyleandelene er av essensiell betydning og inndelingens kategorier er få og sterkt ordnet (sortert). Som regel er et gruppert søylediagram bedre enn et stablet mtp. sammenligninger av kategorier og andeler.

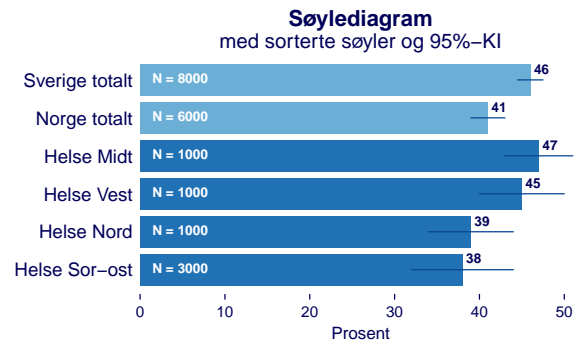
5.2 Mal for linjediagram

Offentliggjøringsmalen for **linjediagrammet** er illustrert i Figur 6 med tilhørende R-kode i appendikset. Grunnprinsippene for søylediagrammet, dvs. prinsippene illustrert av forskjellene mellom Figur 7 og 3, gjelder også for linjediagrammet.

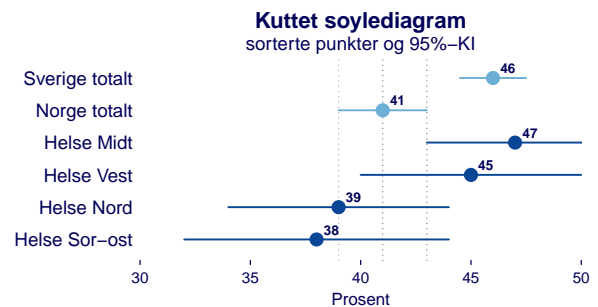
Ved illustrasjon av ett objekts, eller sammenligning av flere objekts, utvikling over tid skal **horisontal akse** angi tiden. Dersom utviklingen ikke foregår langs tidsdimensjonen må den horisontale aksene representere en numerisk variabel med datapunkter som er sterkt sekvensielle, dvs. at de har klart definert rekkefølge. For en generell kategorisk variabel vil ikke sammenhengen mellom påfølgende datapunkt være sterk nok til å rettferdiggjøre linjer mellom punktene. Likeledes er det helt feil å trekke linjer mellom kategorier, f.eks. mellom de individuelle datapunktene i er vertikal versjon av Figur 3.

Den **vertikale aksene** kuttes/tilpasses etter behov mtp. formålet (fremheve forskjeller, absoluttnivåer, trender eller lignende), men den må ha en klart definert retning, dvs. at den må være numerisk (eller kategorisk og sterkt ordnet).

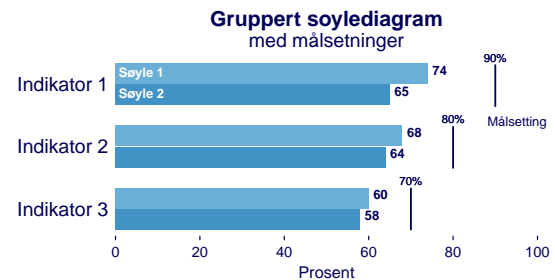
En linje kan f.eks. benyttes til å illustrere overlevelse ved ett sykehus fra år til år, men ikke til å illustrere overlevelse ved ulike sykehus i ett spesifikt år. For å sammenligne årlige verdier flere sykehus benyttes **flere linjer** i samme diagram, men mtp. lesbarhet bør det ikke anvendes mer enn fire linjer per plott, f.eks. to i blånyanser og to i grånyanser.



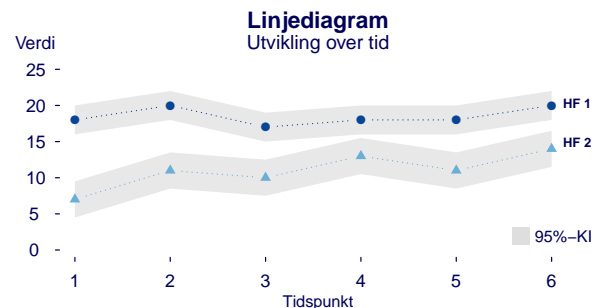
Figur 3. Et enkelt søylediagram.



Figur 4. Et kuttet «søylediagram».



Figur 5. Et gruppert søylediagram.



Figur 6. Et enkelt linjediagram.

Med mindre de sekvensielle datapunktene ligger svært tett, er det, som i **Figur 6**, greit å benytte punktmarkeringer i tillegg til linjer. Dersom et datapunkt mangler skal det ikke trekkes en linje mellom nabopunktene til det manglende punktet. I levetidskurver benyttes sjeldent punktmarkeringer, og for andre spesialfigurer er dette en skjønnsvurdering mtp. lesbarhet.

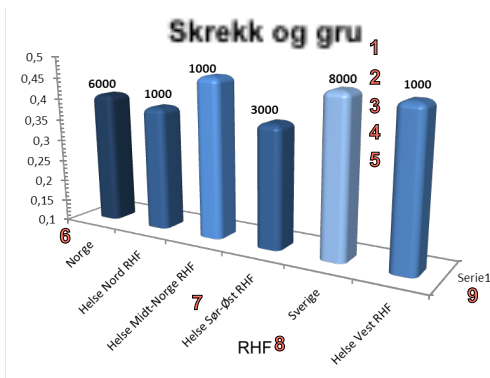
5.3 Mal for andre figurer

Spredningsplottet kan håndteres som et spesialtilfelle av linjediagrammet ved at linjene mellom punktene droppes.

Boksplottet kan likeledes håndteres som et spesialtilfelle av linjediagrammet ved å droppe linjene og erstatte punktene med bokser som angir minimum- og maksimumsverdi, nedre og øvre kvartil, median og eventuelle ekstremverdier.

For **andre figurer** må den figuransvarlige forsøke å legge seg så nær de beskrevne malene som mulig. Levetidskurver kan f.eks. behandles som et linjediagram hvor punktmarkeringer droppes. Som en pekepinne på hvilke aspekter som er viktige under figurutvikling gir **Figur 7** et eksempel på hvordan det *ikke* skal gjøres. En sammenligning av dette skrekkeksempellet med malen i **Figur 3**, som gjengir samme data, gir noen **grunnprinsipper** som er overførbare til andre figurer.

Unngå kakediagram og alle former for 3D- og arealeffekter, disse kompliserer kun avlesning og tolkning av figurene. Et søylediagram er alltid å foretrekke da det kun krever avlesning og sammenligning i en dimensjon (lengderetningen).



Figur 7. Et skrekkeksempel med data fra **Figur 3**. Feilene er: **1)** dårlig oppløsning, **2)** N kan tolkes som verdier, **3)** 3D-effekt, **4)** meningsløse fargeforskjeller, **5)** ingen søylesortering, **6)** kryptisk og kuttet akse, **7)** dårlig lesbarhet og arealbruk, **8)** overflødig/misvisende, og **9)** meningsløst.

5.4 Mal for tabeller

Tabeller skal være minimalistisk og lette å lese, som **Tabell 1**. Unngå alle effekter som fyll, farge, fete skrifttyper, rutenett, rammer og vertikale linjer. Tallkolonner skal være venstrestilt, inkludert kolonneoverskriften, mens kolonner med tall skal være høyrestilt. Numeriske verdier i samme kolonne oppgis med samme presisjon (tenk antall siffer etter komma i desimaltall), men ikke med en høyere presisjon enn nødvendig. Institusjoners dekningsgrader trenger f.eks. ikke å oppgis med desimaltall.

6. Oppsummering

Fargepaletten er primært blå som definert i **Tabell 2** på side 5. Ønsket **filformat** er PDF eller PNG med høy oppløsning.

For numeriske verdier på tvers av kategorier anvendes malen for **søylediagram**, som illustrert i **Figur 3**, eller variasjoner av denne (som f.eks. **Figur 5**). Dersom akse kuttet erstattes søylene av punkter som illustrert i **Figur 4**. For utvikling over tid anvendes malen for **linjediagram**, som er illustrert i **Figur 6**. Overskrifter samt figur- og tabelltekst skal **ikke** inngå i selve grafikkfilene.

For andre figurer anvendes en tilsvarende grafisk profil sammen med visse grunnprinsipper. Veiledning for **figurvalg** er gitt i **Tabell 3** på side 5. Du kan teste din forståelse for grunnprinsippene på www.kvalitetsregistre.no/standardisering-av-resultater-til-offentliggjoring.

Tabeller skal være minimaltisk, som illustrert av **Tabell 1**. En mer detaljert beskrivelse finnes på nettsiden nevnt over.

Bruk følgende punkter som en **sjekkliste**:

1. Vær bevisst på hva data betyr og budskapet du ønsker å kommunisere. Det skal formidles det på en klar, effektiv og presis måte i samsvar med mottakerens forventninger og evner.
2. Vurder om budskapet best formidles av en tabell eller en figur. Avhengig av hva som illustreres velges figurtype fra **Tabell 3** på side 5. Benytt de definerte malene.
3. Velg informative og hensiktsmessige overskrifter, markeringer, etiketter og akseinndelinger.
4. Vurder om enkeltfasetter av figuren skal fremheves spesielt vha. farge, form eller størrelse. Stripp ellers figuren for alt som ikke er vesentlig og gi de gjenværende nødvendighetene en relativt dempet profil.
5. Figuren bør være selvforklarende, eller i det minste vanskelig å feiltolke; la gjerne en utenforstående lekmann studere figuren og fortelle hva han eller hun får ut av den.












Noter

1. www.no.wikipedia.org/wiki/Vektorgrafikk
2. www.en.wikipedia.org/wiki/Pixel_density
3. www.no.wikipedia.org/wiki/Webfarger
4. www.en.wikipedia.org/wiki/Chart_junk













Appendiks

Oversiktstabeller for figurvalg samt link til nettside med andre relevante tema (som formidling av usikkerhet) og eksempel på **R-kode** følger. Koden er ikke hugget i stein, det vil i mange tilfeller være nødvendig med tilpasninger og skreddersøm. De som ikke anvender R får gjenskape layouten skissert i dette dokumentet så godt som mulig.

Farger og figurer

Farge											
Rødt	198	107	066	033	008	000	255	077	115	166	204
Grønt	219	174	146	113	069	000	114	077	115	166	204
Blått	239	214	198	181	148	089	096	077	115	166	204
HEX	#c6dbef	#6baed6	#4292c6	#2171b5	#084594	#000059	#FF7260	#4D4D4D	#737373	#A6A6A6	#DADADA

Tabell 2. RGB-/HEX-koder for offentliggjøringens fargepalett. Maksimumsverdi og metning for RGB-kodene er 255.

Illustreres	Figurtype	Alternativ
Numeriske verdier på tvers av kategorier (Sammenligninger)	 Søylediagram	Med eller uten konfidensintervaller Stigende eller fallende sortering
		 Punktdiagram
Tidsserier (Sekvensielle data)	 Linjediagram	Med eller uten konfidensintervaller
		 Linjediagram
Andeler	 Grp. søylediag.	En gruppe til hvert objekts andeler
		 Stab. søylediag.
Avvik	 Søylediagram	Kan grupperes
		 Søylediagram
Variasjon (Fordelinger)	 Histogram	Kan erstattes med fordelingskurve
		 Boksplott
Assosiasjon	 Spredningsplott	Evt. linje for å assosiasjonens styrke
		 Bobleplott

Tabell 3. Oversikt over hvilke figurer som skal velges avhengig av illustrasjonens formål.

R-kode

R-kode for eksemplene i dokumentet er tilgjengelig på:
www.kvalitetsregistre.no/standardisering-av-resultater-til-offentliggjøring