

Validering av registerdata

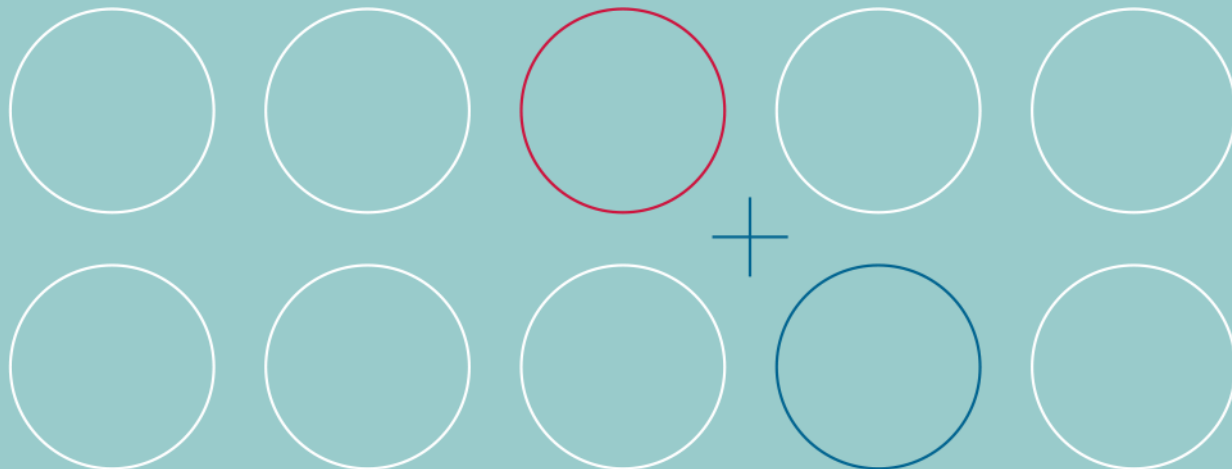
Torunn Varmdal

Forsker, Norsk hjerneslagregister, St. Olavs hospital

Rådgiver, Nasjonalt servicemiljø for medisinske kvalitetsregistre, Midt-Norge

Forsker, Institutt for sirkulasjon og bildediagnostikk, NTNU

Tromsø, 29. mars 2023



Agenda

- Introduksjon til datakvalitet
- Definisjon, hvorfor er datakvalitet viktig
- De seks datakvalitetsdimensjonene
 - Relevans
 - Kompletthet
 - Korrekthet
 - Reliabilitet
 - Sammenlignbarhet
 - Aktualitet

En utfordring/påstand:

Det er lite bevisst forhold til datakvalitet blant forskere og andre brukere av data (?)

Risiko:

- Feilaktige konklusjoner i forskning
- Styrer helsetjenesten på svakt/utilstrekkelig datagrunnlag

Definisjon av datakvalitet i et register:

- Alle trekk og karakteristika ved datasettet som bestemmer dataenes evne til å oppfylle registerets formål.



ISO 8402:1986: Quality - Vocabulary

Definisjon av datakvalitet i et register:

- Alle trekk og karakteristika ved datasettet som bestemmer dataenes evne til å oppfylle registerets **formål**.



ISO 8402:1986: Quality - Vocabulary

Definisjon av datakvalitet i et register:

- Alle trekk og karakteristika ved datasettet som påvirker dataenes evne til å oppfylle registerets **formål**.
- Datakvaliteten defineres således av **brukeren** – data må være egnet til den ønskede bruken.

Klinikeren

- Kvalitetsindikatorer/innhold som beskriver kvalitet på behandlingen
- Relevante medisinske data
- Tilgang til oppdaterte data

Ledelse

- Innhold som beskriver variasjon i kvalitet/behandling
- Tilgang til resultater i egnet format
- Korrekte og komplette data

Forskeren

- Korrekte/komplette data
- Sammenlignbare data (kobling)
- Metadata
- Tilgang til kvalitetssikrede årsdatasett i egnet format

← Eksempler, ikke utfyllende

Data kan sies å være av god kvalitet når relevante, korrekte og komplette data er tilgjengelige for brukerne når de trenger det og i et format de kan benytte.

Hvordan undersøke,
måle og forbedre
datakvalitet i et
register?

Hvordan måle datakvalitet?

- Datakvalitet kan operasjonaliseres gjennom et sett attributter – datakvalitetsdimensjoner – som beskriver ulike aspekter som er viktige for bruken av data
- Ved å definere dimensjonene kan man måle datakvalitet langs alle relevante akser.

Seks datakvalitetsdimensjoner

1. Relevans
2. Kompletthet
3. Korrekthet (validitet)
4. Reliabilitet
5. Sammenlignbarhet
6. Aktualitet

Datakvalitet i kvalitetsregistre:

www.kvalitetsregistre.no/datakvalitet

«Det ideelle kvalitetsregisteret»



Registerets innhold
Relevante data
Godt definerte variabler



Alle enheter og pasienter
Hele populasjonen,
i hht inklusjonskriteriene



Registreringene er:
Komplette
Korrekte
I tide



Data ut
Data tilgjengeliggjort i hht registerets formål

- I tide
- På egnet format
- Dokumentert
- Tolkbart

Hvordan måle datakvalitet?

- Det eksisterer ingen «best practice» eller evidensbasert beste måte å evaluere datakvalitet i helseregistre på.
- En review* av datakvalitetsstudier i register fant tvetydige og/eller mangelfulle definisjoner av kvalitetsattributter i de ulike studiene, og at de samme attributtene ble kalt ulike ting i ulike studier (men de to mest siterte er accuracy og completeness).
- Mange ulike kvalitetsdimensjoner og rammeverk er i bruk.

*Arts et al.: Defining and Improving Data Quality in Medical Registries: A Literature Review, Case Study, and Generic Framework



Land	Datakvalitetsdimensjon	Definisjon
Sverige	Aktualitet	Tid fra hendelsen inntreffer til den er registrert i kvalitetsregisteret
	Kompletthet	Dekningsgrad. Andel av målpopulasjonen som finnes i registeret
	Validitet	I hvilken grad er data og funn sanne for målpopulasjonen (ekstern validitet)
	Sammenlignbarhet	At inklusjonskriterier, registrerings- og kodingsrutiner er tydelige og enhetlige
England	Tilslutning (coverage)	Andel enheter som leverer data av de som skulle ha levert
	Kompletthet	Data har alle enheter som trengs for å måle det man ønsker å måle (på variabelnivå)
	Validitet	I hvilken grad inneholder data lovlige verdier og riktig format for hvert datapunkt
	Defaultverdier	I hvilken grad er default verdier som er spesifisert i standarder brukt
	Integritet	I hvilken grad det er intern konsistens i dataene som er registrert (eks at kjønn samsvarer med diagnose, at samme pasient i ulike datasett har samme fnr)
	Aktualitet	Tiden mellom data blir registrert og leveranse av «produktet» til brukere av data
Canada	Relevans	I hvilken grad oppfyller registeret nåværende og fremtidige behov hos brukerne. Hvor verdifulle er dataene? Kan de endres ved behov?
	Aktualitet	Tid fra hendelsen inntraff til informasjon om hendelsen er tilgjengelig for brukere av data.

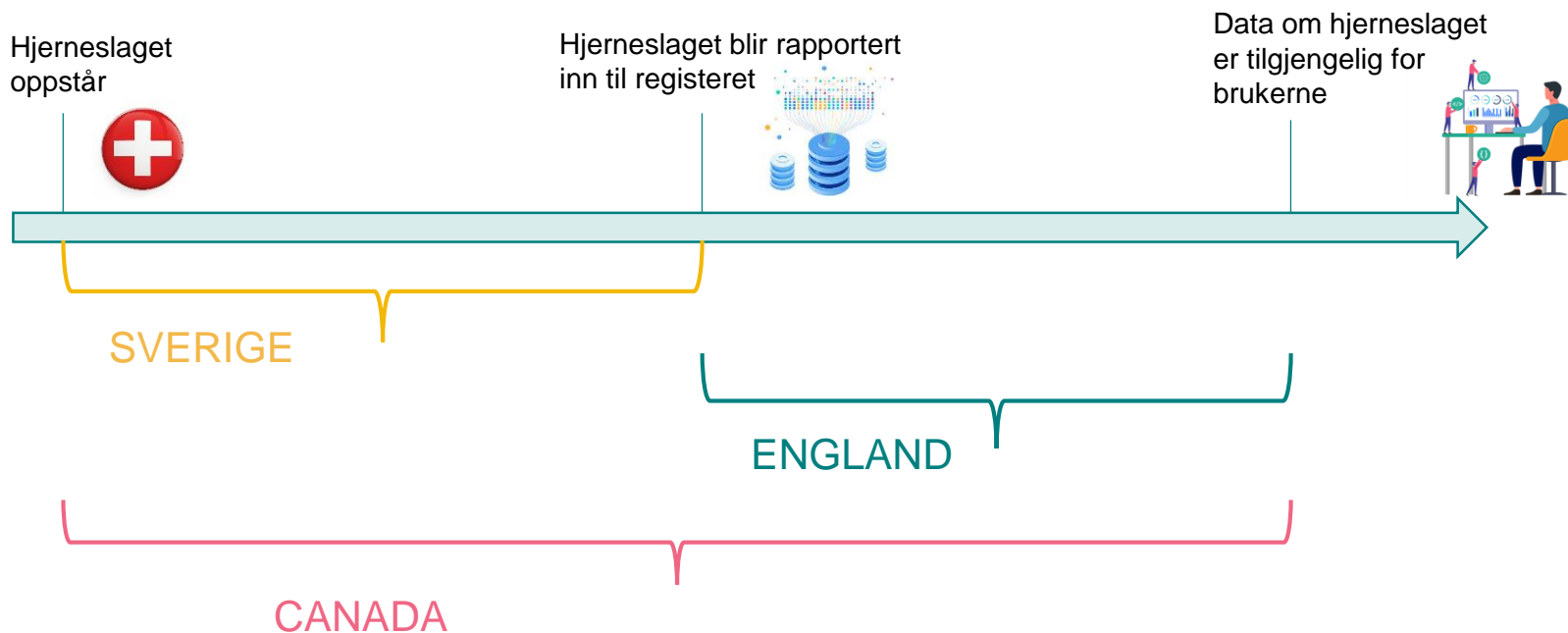


Land	Datakvalitets-dimensjon	Definisjon
Canada forts.	Brukervennlighet (usability)	Om data er lett tilgjengelige og tolkbare
	Korrekthet (Accuracy)	I hvor stor grad data reflekterer virkeligheten den skal måle (intern validitet)
	Sammenlignbarhet	I hvilken grad databasen er konsistent over tid og registerets bruk av standard kodeverk, slik at de blir sammenlignbare (mulige å koble) med andre databaser
EU Parent	Korrekthet (Accuracy)	I hvor stor grad data reflekterer virkeligheten den skal måle (intern validitet)
	Kompletthet	I hvilken grad er alle nødvendige data registrert som burde ha vært registrert. Dekningsgrad, men også kompletthet på variabelnivå.
	Tolkbarhet og tilgjengelighet (Interpretability and accessibility)	Om data er lett tilgjengelige og tolkbare. Om data er tilgjengelige i ønsket/passende format, om det finnes metadata og informasjon om datakvalitet lett tilgjengelig
	Relevans	I hvilken grad data oppfyller registeret nåværende og fremtidige behov hos brukerne. Hvor verdifulle er dataene? Kan de endres ved behov?
	Aktualitet	Tid fra hendelsen inntreffer til informasjon om hendelsen er tilgjengelig for brukere av data.
	Sammenheng (Coherence)	Sammenlignbarhet over tid og på tvers av ulike datakilder (mulighet til å koble data)



Land	Datakvalitetsdimensjon	Definisjon
Sverige	Aktualitet	Tid fra hendelsen inntreffer til den er registrert i kvalitetsregisteret
	Kompletthet	Dekningsgrad. Andel av målpopulasjonen som finnes i registeret
	Validitet	I hvilken grad er data og funn sanne for målpopulasjonen (ekstern validitet)
	Sammenlignbarhet	At inklusjonskriterier, registrerings- og kodingsrutiner er tydelige og enhetlige
England	Tilslutning (coverage)	Andel enheter som leverer data av de som skulle ha levert
	Kompletthet	Data har alle enheter som trengs for å måle det man ønsker å måle (på variabelnivå)
	Validitet	I hvilken grad inneholder data lovlige verdier og riktig format for hvert datapunkt
	Default	I hvilken grad er default verdier som er spesifisert i standarder brukt
	Integritet	I hvilken grad det er intern konsistens i dataene som er registrert (eks at kjønn samsvarer med diagnose, at samme pasient i ulike datasett har samme fnr)
	Aktualitet	Tiden mellom data blir registrert og leveranse av «produktet» til brukere av data
Canada	Relevans	I hvilken grad oppfyller registeret nåværende og fremtidige behov hos brukerne. Hvor verdifulle er dataene? Kan de endres ved behov?
	Aktualitet	Tid fra hendelsen inntreffer til informasjon om hendelsen er tilgjengelig for brukere av data.

Aktualitet – 3 definisjoner



Seks datakvalitetsdimensjoner

1. Relevans
2. **Kompletthet**
3. **Korrekthet (validitet)**
4. Reliabilitet
5. Sammenlignbarhet
6. Aktualitet

Begrepsavklaring

Validering \neq validitet!

Validering: Prosessen med å undersøke og forbedre datakvalitet

Validitet: En datakvalitetsdimensjon (korrekthet)

1. Relevans

I hvor stor grad oppfyller registeret nåværende og fremtidige behov hos brukerne

1. Tilgjengeliggjøring

- All kjent bruk av data fra registeret er dokumentert.

2. Tilpasningsdyktighet

- Registeret evner å endre seg etter behov.

3. Nytteverdi

- Informasjon fra registeret fyller et kunnskapshull, og brukere av data får det de har behov for.

Hvordan validere relevans?

- Dokumentanalyse
 - ✓ Dokumentere nå-situasjonen: i hvilken grad er registeret *tilgjengelig, tilpasningsdyktig* og har *nytteverdi*.
- Brukerundersøkelser
 - Regelmessig undersøkelse av nivået på bruk av data
 - Brukerundersøkelser for å sjekke brukeres tilfredshet
- Utforskende metoder
 - Vurdere datasettet i lys av faglige retningslinjer
 - Benytte journalgjennomgang eller andre registre (NPR, DÅR...) til å belyse ev mangler ved eget register
 - Har registeret overflødige variabler som i liten grad har nytteverdi?

2. Kompletthet

I hvilken grad er alle data registrert som burde vært registrert



Dekningsgrad

- I hvor stor grad er alle tilfellene i målpopulasjonen inkludert i registeret



Mangler pasient

Variabelkompletthet

- I hvor stor grad er registreringene fullstendige



Mangler informasjon om pasient

3. Korrekthet

Andelen tilfeller i registeret med en gitt karakteristikk som faktisk har den aktuelle attributten

Inklusjonskriterier

- I hvor stor grad skulle alle tilfellene i registeret faktisk vært der



Feil pasient

Variabler

- I hvor stor grad er registreringene korrekte



Feil informasjon om pasient

Sammenheng mellom komplettthet og korrekthet*

	GULLSTANDARD			
		Ja	Nei	Total
REGISTERET	Ja	858	12	870
	Nei	107	4215	4322
	Total	965	4227	5192

→ **Korrekthet**
 Positiv prediktiv verdi
 $858/870 = 98.6\%$

↓
Kompletthet
 Sensitivitet
 $858/965 = 88.9\%$

For å vite om registeret vårt fanger målpopulasjonen, må vi vite BÅDE korrekthet og komplettthet!

Obs kompletthet

- Missing completely at random (MCAR)* eller systematisk skjevhet? Noen studier har funnet at det ikke nødvendigvis er tilfeldig hvilke tilfeller som ikke blir rapportert inn:
 - Aboa-Eboule C, Mengue D, Benzenine E, Hommer M, Giroud M, Bejot Y, et al. How accurate is the reporting of stroke in hospital discharge data? A pilot validation study using a population-based stroke registry as control. *Journal of neurology* 2013;260(2):605-13.
 - Ferreira-Gonzales I, Marsal JR, Mitjavila F, Parada A, Ribera A, Cascant P, et al. Patient Registries of acute coronary syndrome: assessing or biasing the clinical real world data? *Circulation Cardiovascular quality and outcomes*. 2009;2(6): 540-7.

*<https://www.ncbi.nlm.nih.gov/books/NBK493614/>

Hvordan måle kompletthet?

Kvantitative (direkte) metoder:

- Sammenligne registerdata mot kildedata eller annet (komplett og korrekt) register
- Capture-recapture metode: estimat av totalpopulasjonen basert på kobling av flere uavhengige datasett

Semi-kvantitative (indirekte) metoder:

- Systematisk monitorering for å oppdage uventede trender i registreringene
- Sammenligne registeret med etablert kunnskap om forventede insidensrater

Hvordan måle korrekthet?

Stemmer registerinnholdet med «sannheten»?

- Sammenligne registerdata mot kildedata. Dette forutsetter at kildedata (journal) er definert som korrekt/sann
- Etablere en gullstandard basert på en uavhengig review og vurdering av kildedata, og sammenligne med registerdata
- Problematisk når det ikke finnes «sann» kildedata, f.eks. PROM/PREM, eller om man ikke får tilgang til kildedata

Eksempel

- Hentet fra studie om korrekthet og kompletthet i Norsk hjerneslagregister og slagdiagnoser i Norsk pasientregister

Varmdal T, Bakken IJ, Janszky I, Wethal T, Ellekjær H, Rohweder G, Fjærtøft H, Ebbing M, Bønnaa KH. Comparison of the validity of stroke diagnoses in a medical quality register and an administrative health register. *Scand J Public Health*. 2016 Mar;44(2):143-9. doi: 10.1177/1403494815621641. Epub 2015 Dec 11.

Metodiske vurderinger og valg

- Problemstilling: Hvor korrekt og komplett er Norsk hjerneslagregister og hjerneslagdiagnoser i NPR?

1. Hvor stor sample?

Alt 1: Beregne utvalgsstørrelse og trekke random sample

Alt 2: Pragmatisk tilnærming basert på ressurser til rådighet, og «bakveisregning»

Beregne utvalgsstørrelse

		Usikkerhet (halv lengde av 95%-KI)		
		± 1%	± 5%	± 10%
Prevalens	1%	380		
	5%	1825	73	
	10%	3457	138	35
	15%	4898	196	49
	20%	6146	246	61

Se tabeller for beregning av utvalgsstørrelse basert på antatt prevalens av feil og ønsket usikkerhetsmargin.

Dette er mer detaljert beskrevet her:

<https://www.kvalitetsregistre.no/sites/default/files/validering-utregning.pdf>

Metodiske vurderinger og valg

- Vi inkluderte totalt 5192 pasienter ved 4 sykehus – fordelt slik:

Tilstand og ICD-10 kode	Registrert med hoveddiagnose	Registrert med bidiagnose
Hjerneslag (I61, I63, I64)	Alle	Alle
TIA (G45)	Alle	25%
Slagsekvele eller annen cerebrovaskulær sykdom (I62, I65-69)	Alle	25%
Epilepsi (G40), migrene (G43), hodepine (G44), vertigo (R42), delirium (F05), hjernesvulster (D33, D43, C71)	Alle	25%
Tilfeldig uttrekk av pasienter uten nevrologiske diagnoser	3.45%	

Metode

- Vi inkluderte totalt 5192 pasienter ved 4 sykehus – fordelt slik:

Tilstand og ICD-10 kode	Registrert med hoveddiagnose	Registrert med bidiagnose
Hjerneslag (I61, I63, I64)	Alle	Alle
TIA (G45)	Alle	25% Vekt x 4
Slagsekvele eller annen cerebrovaskulær sykdom (I62, I65-69)	Alle	25% Vekt x 4
Epilepsi (G40), migrene (G43), hodepine (G44), vertigo (R42), delirium (F05), hjernesvulster (D33, D43, C71)	Alle	25% Vekt x 4
Tilfeldig uttrekk av pasienter uten nevrologiske diagnoser	3.45%	Vekt x 28.97

Metode

- Blindet gjennomgang av journaler, kategoriserte tilfeller i tre grupper: sannsynlig slag, mulig slag, ikke slag
- Alle i kategorien “mulig slag” ble gjennomgått på nytt av lege
- Beregnet sensitivitet, spesifisitet, positiv og negativ prediktiv verdi med 95% konfidensintervall
- Definerte sensitivitet som kompletthet, og positiv prediktiv verdi som korrekthet

Protokoll for gjennomgang av journaler og etablering av gullstandard

Scenarios	A	B	C	D	E
Findings					
Positive CT/MRI *	Yes	No	Yes	No	No
Acute onset of symptoms	Yes	Yes	Yes/No	Yes	Yes/No
Symptoms lasting >24 hours	Yes	Yes	Yes	Yes	No
Conclusive focal symptoms <i>Hemiparesis of arm/face/leg</i> <i>Aphasia</i> <i>Dysarthria</i> <i>Hemianopia</i>	Yes	Yes	No	No	No
Uncertain focal symptoms <i>Tactile numbness/tingling</i> <i>Loss of balance or coordination</i> <i>Swallowing problems</i> <i>Diplopia</i>	Yes/No	Yes/No	Yes/No	Yes	Yes/No
Conclusion	Probable stroke			Possible stroke	No stroke

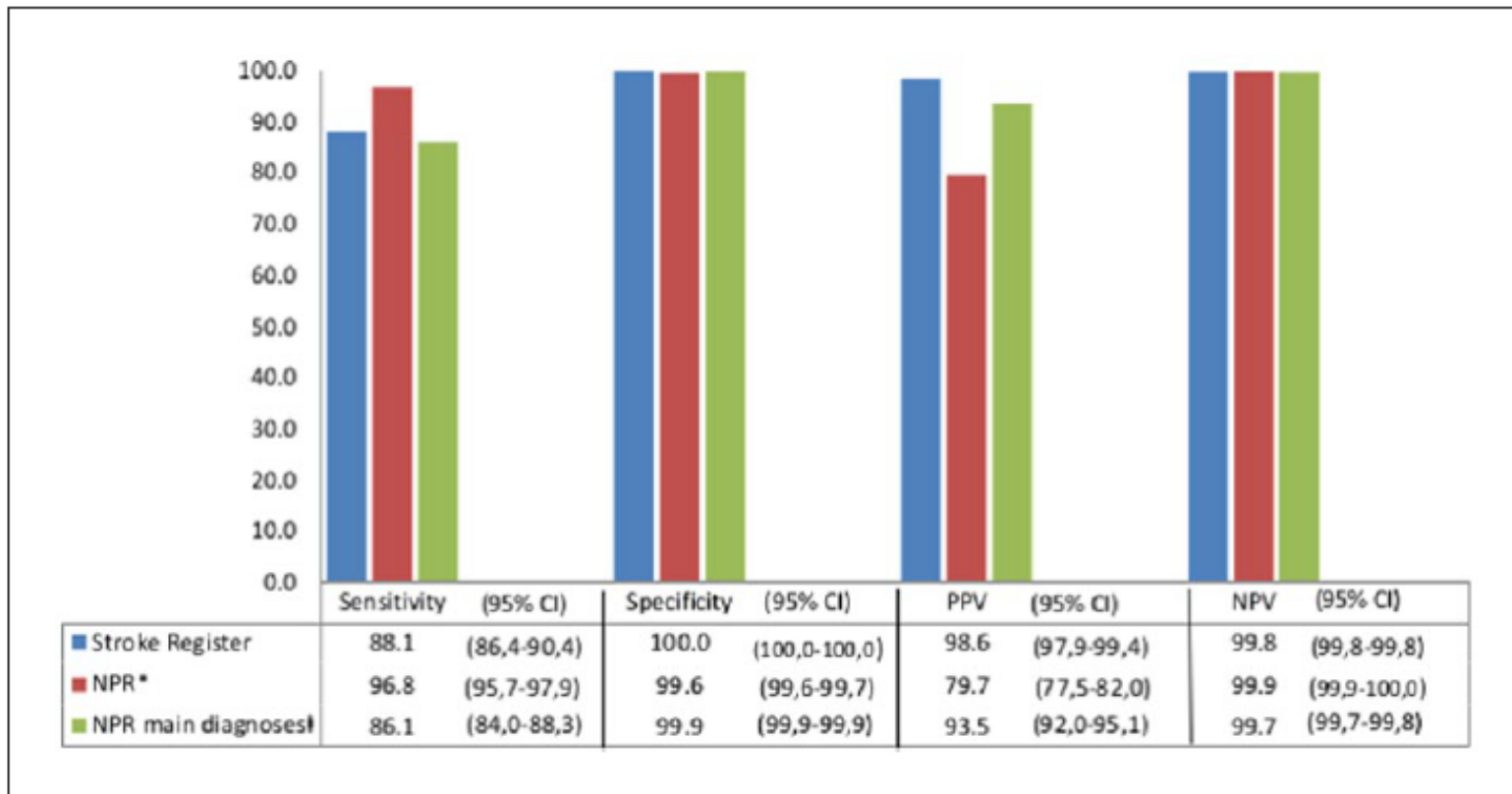
* CT: Computed tomography, MRI: Magnetic resonance imaging.

Analysér

The gold standard

		Yes	No	
Present in the register	Yes	True positive (TP)	False positive (FP)	$TP/(TP+FP)=$ positive predictive value
	No	False negative (FN)	True negative (TN)	$TN/(FN+TN)=$ negative predictive value
		$TP/(TP+FN)=$ sensitivity	$TN/(FP+TN)=$ specificity	

Resultat



Andre funn

- De falske positive i NPR var hovedsakelig feilkoding pga at pas hadde hatt hjerneslag tidligere, eller at pas var inne til rehabilitering
- De falske negative i Norsk hjerneslagregister (de som manglet, men som skulle vært med) var i større grad pasienter som døde på sykehus → tyder på en skjevhet i registeret

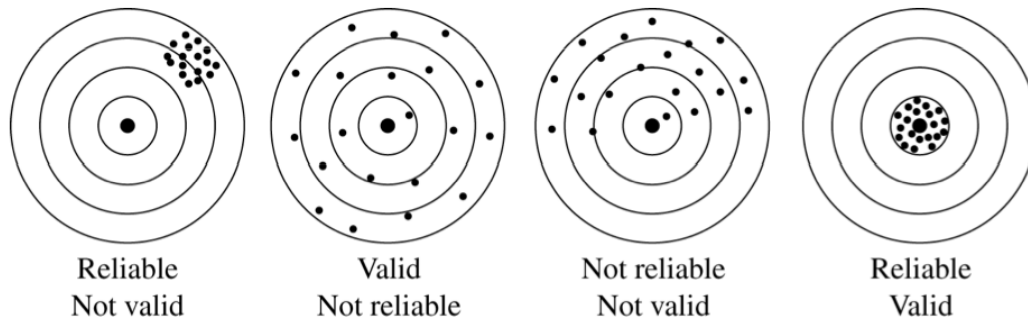
Råd og tips

- Både kompletthet og korrekthet er nødvendig å undersøke for å se om registeret treffer sin målpopulasjon
- Studier kan gjøres på langt færre pasienter enn i vårt eksempel!
- Korrekthet og kompletthet kan undersøkes på lignende måte også på variabelnivå, men da med litt andre analysemetoder
- Det er viktig å utarbeide en protokoll for journalgjennomgang før en begynner

4. Reliabilitet

Hvor stabilt/reproduserbart er innholdet i registeret?

- I register: En metode for å teste deltakernes forståelse for og overholdelse av kodingsregler, og for å identifisere problematiske variabler.
- Mindre ressurskrevende/billigere enn å teste korrekthet (validitet)
- Kan brukes når man ikke har en gullstandard (f.eks. PROM-data)



Hvordan måle reliabilitet?

- Inter-rater reliabilitet: Ulike personer registrerer de samme tilfellene
 - La to eller flere registratorer fylle ut registerskjemaet for et utvalg reelle pasienter. Dette krever at registratorene i prosjektperioden får tilgang til journal.
 - Registeret lager et antall caser som ligner reelle journalnotater. To eller flere registratorer fyller deretter ut registerskjemaet for disse casene
- Intra-rater reliabilitet: La samme person registrere samme «ting» til ulik tid.

Hvordan måle reliabilitet?

- Målemetoder som korrigerer for tilfeldig enighet:
 - Cohen's kapp
 - Gwet's AC_1 og AC_2
 - Specific agreement (positivt og negativt samsvar)*
 - ICC (for kontinuerlige variabler)
- Analyse:
 - Hvilke utvalg (pasientgrupper)?
 - Hvilke variabler har lav reliabilitet?
 - Sannsynlig årsak

*Fagerland M, Lydersen S, Laake P. Statistical analysis of contingency tables. Boca Raton, FL: Chapman and Hall/CRC, 2017.

*de Vet HCW, Dikmans RE, Eekhout I. Specific agreement on dichotomous outcomes can be calculated for more than two raters. J Clin Epidemiol 2017; 83: 85 - 9.

*Lydersen S. Positivt og negativt samsvar. Tidsskr Nor Laegeforen. 2018 Mar 19;138(6). Norwegian. doi: 10.4045/tidsskr.17.0963. PMID: 29557156.

Eksempler:

[Inter-rater reliabilitet av Norsk Tonsilleregister](#)

[Inter-rater reliabilitet av Norsk Hjerneslagregister](#)

[Inter-rater reliabilitet av Norsk Hjerteinfarktregister](#)

5. Sammenlignbarhet

I hvor stor grad er data sammenlignbare på tvers av tid, geografi og ulike datakilder

1. **Intern sammenlignbarhet** beskriver i hvilken grad registerets inklusjonskriterier, variabler og kodeverk er konsistente over tid
2. **Ekstern sammenlignbarhet** beskriver i hvilken grad registeret er sammenlignbart (og mulig å koble) med andre datakilder

Hvordan validere sammenlignbarhet?

Intern sammenlignbarhet (dokumentanalyse):

- Historiske endringer:
 - Hvilke år har registeret samlet data fra.
 - Har inklusjonskriteriene har vært konsistente i alle år. Hvis ikke, angi på hvilken måte kriteriene har endret seg og tidspunkt for endring.
 - Har det vært endringer for:
 - Populasjon
 - Innrapporterende enheter
 - Kodeverk og variabler.

Hvordan validere sammenlignbarhet?

Ekstern sammenlignbarhet (dokumentanalyse):

- Koblingsmuligheter:
 - Inneholder registeret en unik person-ID som muliggjør kobling med andre datakilder (f.eks. fødselsnummer).
 - Inneholder registeret en unik enhets-ID som muliggjør identifisering av enheter/organisasjoner ved kobling mot andre datakilder.
 - Inneholder registeret en unik geografisk ID som muliggjør sammenligninger på geografisk nivå ved kobling mot andre datakilder (f.eks. region, fylke, kommune).
 - Er det begrensninger ved kobling mot andre datakilder?

6. Aktualitet

Tid fra en hendelse skjedde til hendelsen er registrert i registeret

ELLER: Tid fra en hendelse skjedde til hendelsen er klar for bruk fra registeret

- Spesielt viktig når man skal bruke data i kvalitetsforbedring, løpende monitorering og styring

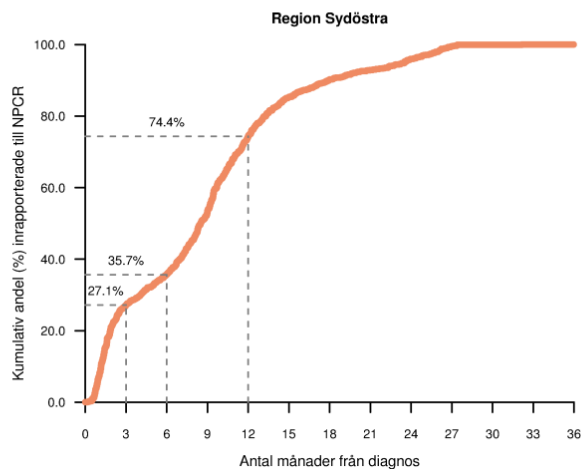
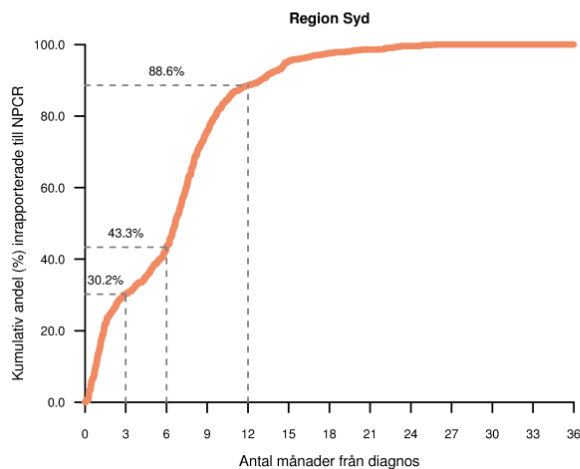
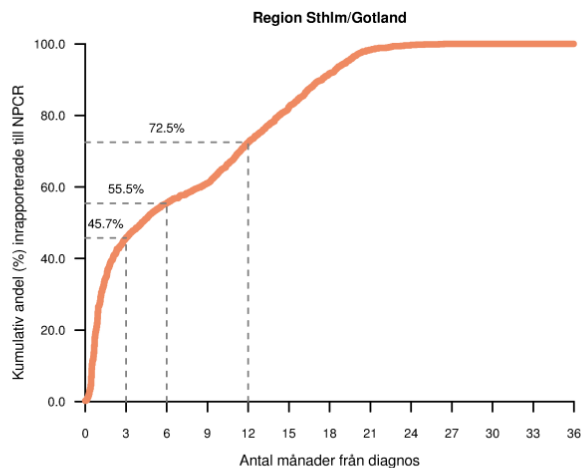
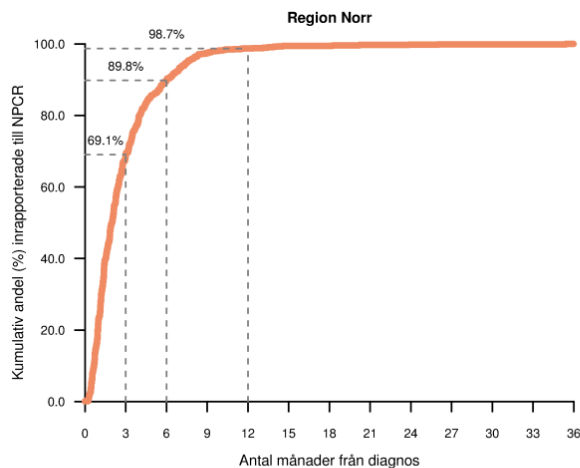
Hvordan måle aktualitet?

Aktualitet oppgis gjerne i absolutte tall per sykehus/innregistrerende enhet. Dette kan angis som:

- Median antall dager
- Gjennomsnittlig antall dager
- Andel hendelser innrapportert innen en gitt tid (f.eks. andel innrapportert innen 30 dager)

Man bør inkludere et mål på spredning; kvartiler hvis man oppgir median, standardavvik hvis man oppgir gjennomsnitt og konfidensintervall hvis man oppgir andeler.

Hvordan måle aktualitet?



Oppsummering

Ved all bruk av registerdata er det viktig å vite noe om datakvaliteten!

- Som bruker av registerdata bør du derfor:
 - Ha et bevisst forhold til datakvaliteten
 - Be registeret om metadata og dokumentasjon på datakvalitet
 - Ev gjøre egne undersøkelser eller studier
- Anbefalt nettside på norsk: [Datakvalitet | Nasjonalt servicemiljø for medisinske kvalitetsregistre](#)