

Datakvalitet i et nasjonalt medisinsk kvalitetsregister

- Hvordan måle andel korrekte registreringer

Ida Almenning Kiel, Ragna Elise Govatsmark, Veronica Berge, Kari Krizak Halle, Kaare Harald Bønna
Seksjon for medisinske kvalitetsregistre, St. Olavs hospital, Trondheim

Bakgrunn

Tidligere datakvalitetsstudier har sett på totalt andel korrekte registreringer per variabel, men lite er gjort for å se på korrekt registrering av variabelenes ulike svaralternativ. Formålet med studien var å undersøke andel korrekte registreringer per svaralternativ for et utvalg av variabler i Norsk hjerteinfarktregister (NHIR), sammenlignet med en gullstandard.

Metode

Syv leger registrerte et tilfeldig utvalg av pasienter i NHIR med ST-segment elevasjon hjerteinfarkt (STEMI) (n=262) og ikke ST-segment elevasjon hjerteinfarkt (NSTEMI) (n=328) på nytt. Legene var blindet for tidligere registrering i NHIR. Andel korrekte registreringer i NHIR, sammenlignet med gullstandard, ble beregnet for hver variabel og separat for alle svaralternativ. For en dikotom variabel tilsvarer dette sensitivitet eller spesifisitet. Kontinuerlige tidspunktvariabler ble kategorisert for å se på andel med og uten avvik i registreringen, og intraklassekorrelasjonskoeffisienten (ICC) ble beregnet.

Eksempel på utregning av andel korrekt klassifiserte registreringer sammenlignet med en gullstandard.

Variabelen viser antall registreringer av «Ja», «Nei», eller «Ukjent» på spørsmål om pasienten utviklet hjertesvikt som komplikasjon under oppholdet.

		Gullstandard			Totalt
		Nei	Ja	Ukjent	
Register	Nei	460	43	0	503
	Ja	46	64	1	111
	Ukjent	3	0	0	3
Totalt		509	107	1	617

Totalt andel korrekte registreringer:

$[(460+64)/617] * 100 = 84,9 \%$

Svaralternativ «Nei»:

$(460/509) * 100 = 90,4 \%$

Svaralternativ «Ja»:

$(64/107) * 100 = 59,8 \%$

Svaralternativ «Ukjent»:

$(0/1) * 100 = 0,0 \%$

Resultat

Eksempel på andel korrekte registreringer for en variabel som i stor grad kan gjenfinnes som en kategorisk variabel i journal

STEMI eller NSTEMI diagnose

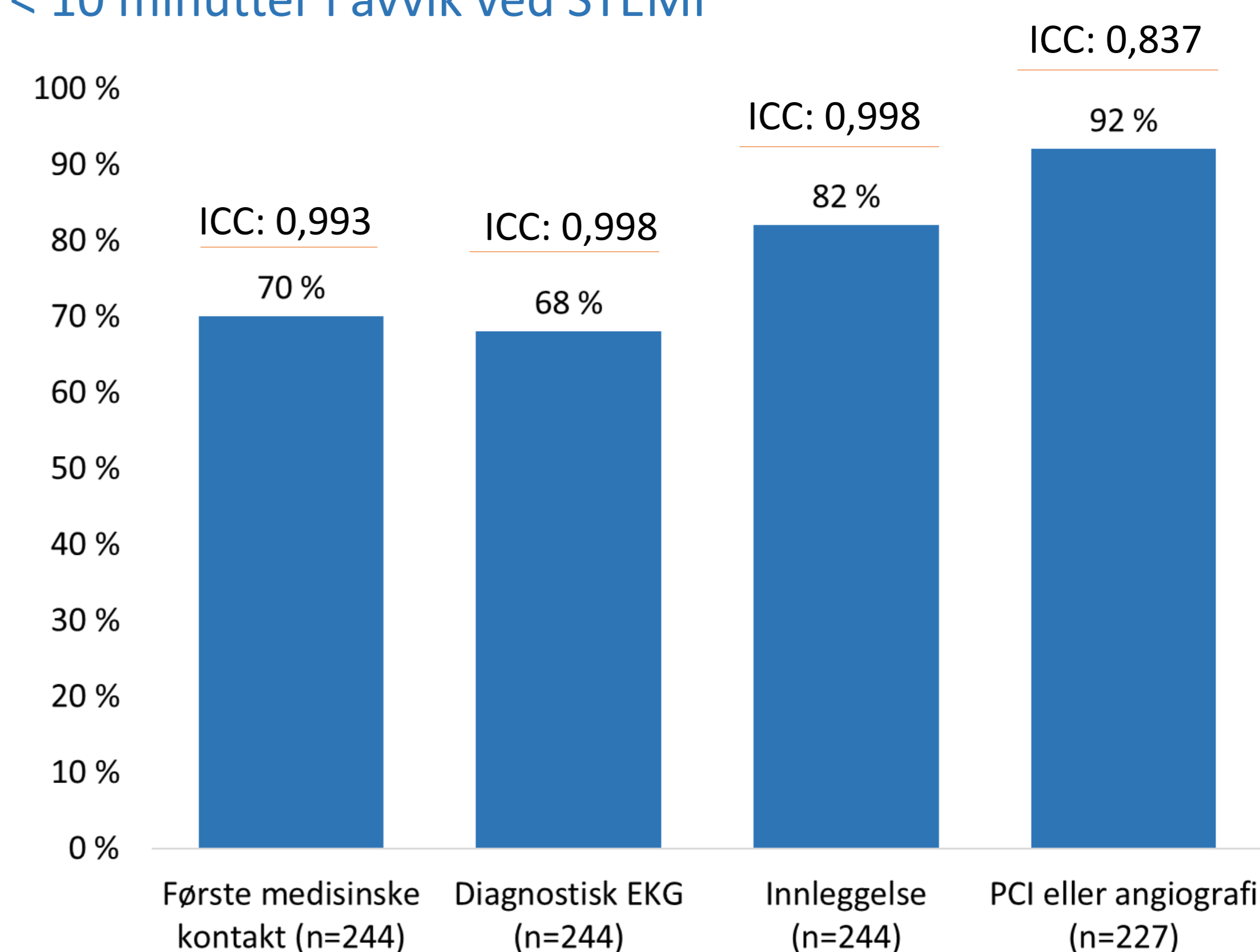


Eksempel på andel korrekte registreringer for en variabel som i stor grad er basert på tolkning av fritekst i pasientjournal

Vedvarende/residiverende/nye brystmerter



Andel registreringer med samme klokkeslett eller < 10 minutter i avvik ved STEMI



Hovedfunn

- Datakvaliteten varierte mellom ulike variabeltyper og mellom ulike svaralternativ
- For kategoriske variabler, som kan gjenfinnes som en kategorisk variabel i journal, var det stort sett høy andel korrekte registreringer per svaralternativ
- For kategoriske variabler, som i stor grad er basert på tolkning av fritekst i pasientjournal, var det lavere andel korrekte registreringer per svaralternativ
- En høy total andel korrekte registreringer for en variabel kan kamuflere en lav andel korrekte registreringer for svaralternativ som er klinisk viktige, men som forekommer sjelden
- På enkelte viktige klokkeslett var andel korrekte registreringer lav, til tross for at ICC var høy

Konklusjon

Andel korrekte registreringer per svaralternativ gir viktig informasjon om en variabel. Beregning av andel korrekte registreringer per svaralternativ, i tillegg til totalt for variabelen, kan benyttes av helseregistre for å undersøke datakvaliteten. ICC er en velkjent metode for å se på korrekthet/samsvar på kontinuerlige variabler, men metoden har begrensninger og var ikke egnet for å undersøke datakvaliteten av kontinuerlige variabler i denne studien.