

Reliabilitet av den norske versjonen av Timed Up and Go (TUG)



Pernille Botolfsen,
fysioterapeut, M.Sc.,
høgskolelektor,
Avdeling for Helsefag,
Fysioterapi, Høgskolen
i Oslo. Epost: pernille.
botolfsen@hf.hio.no

Jorunn L. Helbostad, fysioterapeut, Ph.D.,
forsker, Institutt for nevromedisin, Medisinsk
fakultet, NTNU og Avdeling for geriatri,
St. Olavs Hospital.

Oppgitte interessekonflikter: Ingen.

Denne vitenskapelige originalartikkelen,
innsendt 08.07.09 og akseptert 16.02.10, er
eksternt fagvurdert etter Tidsskriftet Fysio-
terapeutens retningslinjer på www.fysioterapeuten.no og redigert av Kjartan Vårbakken.

Sammendrag

- **Hensikt:** Vurdere intratester-, intertester-, og test-retest reliabilitet av den norske versjonen av «Timed Up and Go» (TUG).
- **Design:** Metodestudie som benytter tverrsnittsstudie-design.
- **Materiale og metode:** Tretti personer (20 kvinner og 10 menn) over 75 år (gjennomsnitt 82,5) gjennomførte TUG tre ganger. To fysioterapeuter skåret deltagerne. For utregning av relativ reliabilitet ble intraclass correlation coefficient (ICC) anvendt, og for utregning av absolutt reliabilitet ble intrasubject standard deviation (S_w) anvendt.
- **Resultater:** Gjennomsnittstiden ble noe lavere ved hver gjennomføring, fra 20,2 til 18,1 sekunder. Intratester- og intertesterreliabilitet målt med ICC(1,1) viste 0,99 for alle målinger. $1,96 S_w$ varierte fra $\pm 0,5$ til $\pm 1,1$ sekund mellom testerne for intratesterreliabilitet, og var $\pm 0,3$ sekunder for intertesterreliabilitet. Alle ICC(1,1) verdiene for test-retest var høyere enn 0,81 og $1,96 S_w$ varierte fra $\pm 3,5$ til $\pm 4,4$ sekunder. Reliabilitet ved gjennomsnittet av to målinger ICC(1,2) var over 0,90.
- **Konklusjon:** Studien viser at den nye norske protokollen av TUG har meget god intratester-, intertester- og test-retest reliabilitet. For høyere relativ reliabilitet anbefaler vi at gjennomsnittet av to målinger benyttes i klinikken. Med endringene foreslått i forhold til instruksjon, tidtakingparametrene og bruk av gjennomsnittet av to målinger, anbefaler vi at den utbedrede norske protokollen tas i bruk i Norge.
- **Nøkkelord:** standardiserte tester, klinisk protokoll, reliabilitet.

Innledning

Teoretisk bakgrunn for valg av studie

Den fysioterapeutiske funksjonsvurderingen var inntil begynnelsen av 1990-årene primært knyttet til den enkelte terapeuts observasjonsevne og kunnskap innen normal funksjon, anatomi og fysiologi. Tolkningen og vurderingen av observasjonene varierte mye og var en kilde til stor subjektivitet (1). I løpet av 1990-årene og frem til i dag har det vært et øket fokus på å utvikle objektive tester for å dokumentere fysioterapeutenes praksis. Flere reliabilitets- og validitetsstudier har vært gjennomført, retningslinjer for

bruk av testene har blitt utarbeidet, og det har vært satsset for å utdanne og støtte fysioterapeuter i bruk av standardiserte tester (2).

Reliabilitet

Reliabilitet dreier seg om dataenes pålitelighet, og sier i hvilken grad testen er reproduserbar eller hvor stor feilmargin metoden er beheftet med. Dersom flere uavhengige målinger av ett og samme fenomen gir samme eller tilnærmet samme resultat, anses testen som reliabel (1). Tester som benyttes i klinisk praksis må være reliable innen testere (intratester) og mellom testere (intertester),

samt mellom forskjellige observasjoner (test-retest). Jo høyere reliabiliteten er, jo lettere er det å påvise reelle forskjeller og sammenhenger mellom uavhengige målinger. Det er i hovedsak tre faktorer som kan påvirke målingene og som kan redusere reliabiliteten: testeren som utfører målingene, selve testen eller måleinstrumentet og personen som undersøkes (1). Noen teorier om reliabilitet betrakter alle avvik som målefeil, mens andre skiller biologisk variabilitet innen individet fra systematiske eller tilfeldige målefeil som skyldes instrument eller testprosedyrer (1).



TUG viser god reliabilitet og er klar for bruk i norske klinikker og forskningsmiljøer.

Utvikling av testen «Timed Up and Go» (TUG)

Det er utviklet flere tester for å si noe om eldres balanse, gange og eventuelle risiko for å falle (3-7). «Get Up and Go» (GUG) testen var utviklet av Mathias og medarbeidere for å vurdere eldres balanse og for å vurdere eldres risiko for å falle (8). «Timed Up and Go» (TUG) ble videre utviklet fra GUG av Podsiadlo og Richardson (3) og måler tiden det tar for personen å reise seg opp fra en 46 cm stol med armlener, gå tre meter, snu, gå tilbake til stolen, snu og sette seg ned. Podsiadlo og Richardson beskriver at testen utføres to ganger, og at første forsøk betraktes som prøverunde, mens andre forsøk er tellende. Personen bruker vanlige sko, og det ganghjelpemiddelet hun eller han vanligvis bruker. Podsiadlo og Richardson rapporterte god intratester- og intertesterreliabilitet av TUG og god samtidig validitet med Bergs balanseskala (3). Forfatterne konkluderte med at TUG var en praktisk, reliabel test for skrøpelige eldres mobilitet. De mente at testen ville kunne være et screeningverktøy i forhold til å si noe om behov for videre testing eller ikke, og en god test for å si noe om personens balanse, ganghastighet eller generelle funksjon (3).

TUGs reliabilitet

Siden 1991 har det vært gjennomført mange reliabilitets- og validitetsstudier av TUG. Noen viser god intratester-, intertester- og test-retest reliabilitet i ulike grupper (3;4;6;9), mens andre har vist lavere reliabilitet (4;5). Definisjonen av intratester- og intertesterreliabilitet har variert (5), intervallet mellom målingene har variert og målingene har vært utført i forskjellige omgivelser (4).

TUGs validitet

TUG har vist god samtidig validitet med den fysiske delen av Sickness Impact Profile, 68 item versjonen (SIP68) (9), med Bergs balanseskala, Barthel Mobility subscale, Tinetti mobility Index (10), selvrapporert gangvanskeindeks (11), ganghastighet (12;13) og med Functional Stair test (14).

En metaanalyse av Bohannon omfattet 21 publiserte studier av TUG i perioden 1990

til 2005, viser stor variasjon i utvalg, antall deltagere, stolhøyde, instruksjon, antall repetisjoner av testen og når tidtaking begynte (15). Slike forskjeller kan forklare ulikheter i resultat fra tidligere reliabilitets- og validitetsstudier av TUG.

TUGs testprosedyre

Bohannons studie viser hvor stor variasjon det kan være for testprosedyrer av en og samme test. Studier utført på TUG i Norge viser tilsvarende variasjon (11;12;16-18). Antall repetisjoner av utførelsen og hvor lang hvilepause personen har hatt mellom målingene har variert i tidligere studier (15). I originalstudien av Podsiadlo og Richardson gjennomførte deltagerne en prøverunde og et tellende forsøk. Det står ikke hvor lang pause deltagerne hadde mellom målingene (3). Andre studier har ikke eksplisitt skrevet hvor mange målinger som er gjennomført (11;16). Noen har benyttet en måling, og det er usikkert om en prøverunde er gjennomført (13;17), hos andre er en måling (14), eller en måling og en prøverunde (12;19;20), eller gjennomsnittet av to målinger benyttet (6;18).

Det er stor variasjon mellom studier i forhold til når tidtaking begynner og slutter. Noen tar tiden fra sete forlater stolen (12;13), andre starter tiden på instruksjonen «gå» (6;17;20), mens noen ikke skriver spesifisert fra når tiden tas (3;11;16;19). Det er tilsvarende variasjon i når tidtakingen stopper.

TUG benyttes i dag i klinisk forskning og i klinikken både i og utenfor Norge. Dr. med. Olav Sletvold oversatte TUG til norsk i 1996 (20), men testen ble den gang ikke reliabilitets- eller validitetstestet, og testprotokollen inneholder få presiseringer med hensyn til gjennomføring.

Hensikt og forskningsspørsmål

Denne studien hadde derfor som mål å undersøke reliabilitet av den norske versjonen av TUG med ytterligere presiseringer av testprotokoll og tidtaking enn i versjonen fra 1996. I tillegg ønsket vi å undersøke hvor mange ganger testen burde gjennomføres for å få et mest mulig reliabelt mål. Studien

var en del av en større reliabilitetsstudie av «Expanded Timed Up and Go» (ETUG) (7). Hensikten ble raffinert i forskningsspørsmålene:

- Hvordan er intratester-, intertester- og test-retest reliabiliteten av den nye norske versjonen av TUG på en gruppe eldre personer over 75 år med fall- og gangproblemer?
- Hvor mange ganger bør TUG gjennomføres for å få et mest mulig reliabelt mål?

I denne studien undersøkte vi intratester-, intertester- og test-retest reliabilitet av TUG. Fokus for intratester- og intertesterreliabilitet var å undersøke målefeil relatert til testen eller testprosedyren, mens for test-retest studien var fokus på å vurdere variasjon i prestasjonen til personen som ble undersøkt (1).

Materiale og metode

Design

Metodestudien benytter en tverrsnittsdesign. Ved undersøkelse av intratesterreliabilitet utførte to testere to målinger på samme observasjonen, i direkte tid og fra video. For intertester-reliabilitet utførte to testere målinger på den samme observasjonen, i direkte tid og fra video (1). I test-retest studien sammenlignet de to testerne tiden på tre forskjellige utførelser av testen av samme person, både i direkte tid og fra video.

Utvalg

I perioden september 07 til juni 08, ble et strategisk utvalg på 30 hjemmeboende personer på 75 år eller eldre fra Oslo og Akershus kommune (20 kvinner og 10 menn) rekruttert. Dette fra Aker universitetssykehus, avdeling for rehabilitering og geriatri. Inklusionskriteriene var at deltagerne: 1) hadde falt minst én gang i løpet av det siste året, eller 2) brukte en eller annen form for ganghjelpemiddel til daglig, og 3) kunne gå minst 10 meter uten ganghjelpemiddel.

Studien ekskluderte personer med demens eller kognitive problemer (Mini Mental Status på < 22) eller som ikke forsto instruksjonen i TUG.

Etikk

Studien er vurdert og godkjent av Regional etisk komité for medisinsk forskningsetikk, Sør-Norge og Norsk samfunnsvitenskapelig datatjeneste. Gjennomføringen av studien følger Helsinki-deklarasjonen som angitt i Helseforskningsloven. Førstegangskontakt med deltagerne ble gjort av ansatte ved Aker universitetssykehus. Deltakelse var frivillig etter skriftlig og muntlig informasjon om studiens innhold og gjennomføring. Deltakerne kunne trekke seg fra studien til enhver tid uten nærmere begrunnelse og uten konsekvenser for den videre oppfølgingen på Aker universitetssykehus.

Beskrivelse av testerne

To fysioterapeuter, tester A og B, med lang erfaring innen rehabilitering og arbeid med eldre og med bruk av TUG, skåret alle deltagerne. En tredje fysioterapeut var ansvarlig for å gi all informasjon og instruksjon til deltagerne, samt ta video av gjennomføringene.

Testprosedyre

Vi tok tiden det tok deltageren å reise seg opp fra en 46 cm stol med armlener (17 cm opp fra stolsete), gå tre meter, snu, gå tilbake til stolen, snu og sette seg ned. Alle deltagerne satt tilbakelent i stolen ved start, og tre-meteren var markert med en 50 cm lang og 1,8 cm bred rød tape på gulvet. Tidtakingen ble startet på instruksjonen «gå» og stoppet da personen satt med sete ned på stolen igjen. En manuell stoppeklokke ble brukt for tidtaking og tiden målt i 1/10 av et sekund. TUG ble gjennomført etter Podsiadlo og Richardsons testprosedyre oversatt av Sletvold (3;20), med unntak av antall gjennomføringer og instruksjon.

Alle deltagerne ble testet i samme rom med noen få minutter pause mellom hver repetisjon av testen. Deltagerne gjennomførte testen tre ganger med tidtaking; en prøverunde og to tellende målinger. Ved behov fulgte en fysioterapeut deltagerne under testingen. I slike tilfeller var det ingen fysisk kontakt mellom terapeut og deltager, og ingen tilleggsinstruksjon ble gitt. Det ble tatt videoopptak av alle observasjonene fra siden med et videokamera tilpasset Mini DV kassetter.

TUGs originale instruksjon er «He is instructed that, on the word «Go» he is to get up and walk at a comfortable and safe pace to a line on the floor 3 meter away, turn, return to the chair, and sit down again» (3).

Instruksjonen til Sletvold (1997) er: «Etter «klar gå» reiser du deg opp og går uten assistanse til merke 3 meter frem på gulvet, snur, går tilbake til stolen og setter deg ned igjen». Vi endret noe på instruksjonen for denne studien etter å ha lest forskjellige varianter i andre studier (5;6;11;12;15-17) til: «Etter «klar gå» så reiser du deg opp, går forbi den røde streken, snur, går tilbake til stolen og setter deg ned igjen». Da deltagerne gikk andre og tredje gang var instruksjonen kun «klar gå» tilsvarer Sletvolds protokoll (20).

I databearbeidingen ble alle tre observasjonene registrert fra videoopptakene. Dette ble utført ved Høgskolen i Oslo. Testerne satt foran en TV og datamaskin da skåringene ble utført. Musen på datamaskinen ble benyttet som en manuell stoppeklokke, og start og stopp ble lagret i en programvare laget for formålet.

Data analyse

Statistiske analyser ble gjennomført i SPSS 16.0. Beregningene for relativ og absolutt reliabilitet følger kriteriene for evaluering av måleinstrumenter for balanse utarbeidet av The Prevention of falls Network Europa (21).

Vurderingsvariabler

Relativ reliabilitet: Intraclass correlation coefficient

Vi benyttet Intraclass correlation coefficient (ICC), modell ICC(1,1) og ICC(1,3), med 95 prosent konfidensintervall som mål på relativ reliabilitet for intratester-, intertester og test-retest reliabilitet (1;22;23). For test-retest reliabilitet var i tillegg ICC(1,2) beregnet for å vise gjennomsnittet av to og to målinger, mellom prøverunden og første måling, og mellom første og andre måling både for tester A og B. ICC beregnes ved hjelp av enveis og toveis variansanalyse (ANOVA) for å estimere varians (en type variabilitet) innen subjekter i forhold til varians mellom subjekter (23). Med ICC(1,1) er systematisk og tilfeldig intrasubjekt-variabilitet sett på som målefeil. ICC(3,1) justerer for middelverdien og er et uttrykk for tilfeldig variabilitet mellom de to målingene (23;24). Det vil si at dersom ICC(1,1) og ICC(3,1) er like er det ingen systematiske feil (25).

Korrelasjonskoeffisienten varierer fra 0 til 1, hvor 1 betyr perfekt sammenheng mellom målingene (1;26). Hva som er en akseptabel korrelasjon er ikke absolutt definert, men Munro foreslår ICC mellom 0,50-0,69

som moderat korrelasjon, mellom 0,70-0,89 som god korrelasjon og 0,90-1,00 som meget god korrelasjon (1;26).

Absolutt reliabilitet: Within-subject standard deviation (S_w)

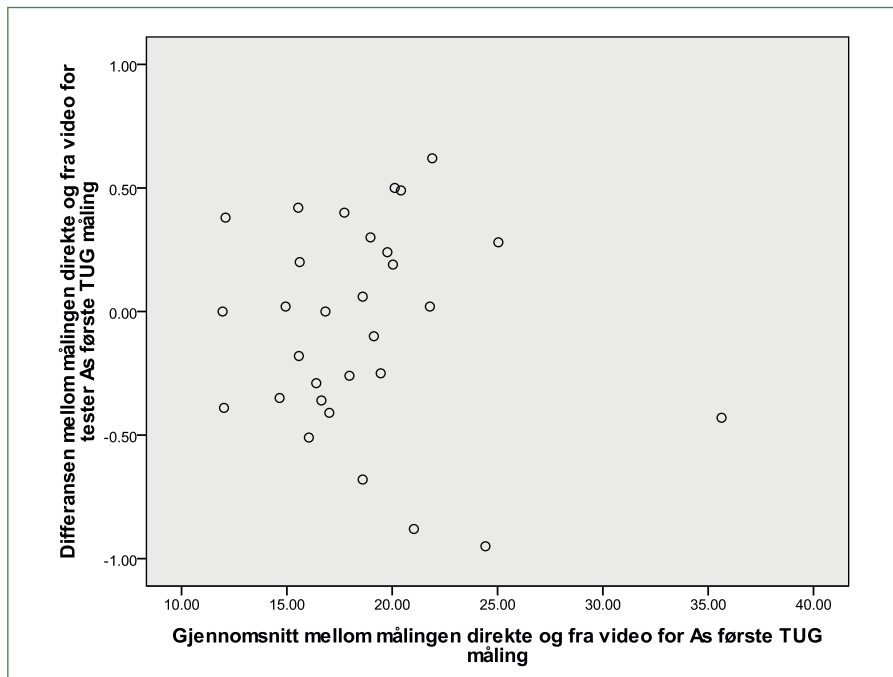
Vi benyttet Within-subject standard deviation (S_w) som oppgis i den aktuelle måleenheten, her sekunder, som uttrykk for absolutt reliabilitet (24;25). Jo høyere S_w , jo større målefeil. For å kunne beregne den absolutte reliabiliteten ved hjelp av S_w må en ha flere deltagere med minst to målinger hver. ANOVA gir oss gjennomsnittet av variansen innen individet. S_w ble beregnet som kvadratrotten av gjennomsnittet av variansen mellom målingene (Z), $S_w = \sqrt{Z}$ (24). Forskjellen mellom en måling og den sanne verdien er forventet å være mindre enn 1,96 S_w for 95 prosent av observasjonene (22). 1,96 S_w er derfor rapportert som mål på intratester-, intertester- og test-retest reliabilitet. Forutsetningene for å bruke S_w er at det ikke er et proporsjonalt forhold mellom variabiliteten og gjennomsnittet av målingene (eksempel figur 1). Kriteriene for bruk av S_w er oppfylt i denne studien.

Resultat

Tretti personer (20 kvinner og 10 menn) fra 75 til 92 år gjennomførte testingen. Tjuefem (83 prosent) hadde falt minst én gang i løpet av det siste året, og 28 (93 prosent) brukte en eller annen form for ganghjelpemiddel til daglig både inne og ute. Andre bakgrunnsvariabler er presentert i tabell 1. Tjuefem (83 prosent) benyttet ett eller to armlener da de reiste seg opp av stolen under testingen. Alle deltagerne benyttet gode sko og alle gikk uten ganghjelpemiddel under testingen.

Reliabilitet

Det var små tidsdifferanser, i absolutt verdier, mellom målingene innen og mellom testerne (tabell 2 og figur 2). Intratester- og intertesterreliabilitet målt med ICC(1,1) var 0,99 for alle målingene. 1,96 S_w varierte fra 0,5 til 1,1 sekunder for intratesterreliabilitet, og var 0,3 sekunder for intertesterreliabilitet (tabell 3 og 4). Alle de relative og absolutt reliabilitetsmålene samsvarte med resultatmålene i absolutt verdier. ICC(1,1) og ICC(3,1) var identiske, som betyr at det ikke var noen systematiske forskjeller mellom målingene. Derfor er kun ICC(1,1) rapportert. I databearbeidingen ble alle tre observasjonene registrert fra både de direkte målingene og



FIGUR 1 Blant og Altman plottdiagram for gjennomsnittet og differansen mellom tester As første og andre måling (målt i sekunder).

TABELL 1 Bakgrunnsvariable (N=30)

	\bar{x}	SD
Alder (år) (min-maks)	82,5 (75 - 92)	(4,42)
Høyde (m) (min-maks)	1,66 (1,47 - 1,80)	(0,08)
Vekt (kg) (min-maks)	64,6 (38 - 93)	(14,8)
Antall forskrevne medikamenter) (min-maks)	6,7 (2 - 11)	(2,4)
Antall medisinske diagnoser) (min-maks)	4,8 (2 - 8)	(1,6)

Medisinske diagnoser	Tidligere		Ved testing	
	N	%	N	%
Hjerneslag	12	(40,0)	12	(40,0)
Hjertelidelser	18	(60,0)	21	(70,0)
Høyt blodtrykk	12	(40,0)	15	(50,0)
Luftveis sykdommer	1	(3,6,7)	12	(40,0)
Diabetes mellitus	7	(23,3)	7	(23,3)
Muskel - skjelett lidelser	15	(50,0)	15	(50,0)
Kognitiv svikt	5	(16,7)	0	(0)
Depresjon	5	(16,7)	0	(0)
Inkontinens	14	(46,7)	14	(46,7)
Nedsatt syn	21	(70,0)	21	(70,0)
Nedsatt hørsel	12	(40,0)	12	(40,0)
Besvimelsestendens	7	(23,3)	5	(16,7)
Andre	26	(86,6)	26	(86,6)

\bar{x} = Gjennomsnitt, min-maks = minimum og maksimum skåren, SD = standard avvik, N = antall

fra videomålingene. ICC-verdiene fra de direkte målingene og fra videomålingene var identiske og resultatene fra videomålingene

er derfor ikke rapportert for intertester- og test-retest reliabiliteten. Intratesterreliabilitet er nettopp målingene mellom de direkte

målingene og videomålingene, så de er rapportert for alle tre observasjonene.

ICC(1,1) for test-retest reliabilitet var høyere enn 0,81 for alle målingene, og det var ubetydelig forskjell i ICC(1,1) mellom prøverunden og første måling i forhold til mellom første og andre måling. Absolutt-reliabiliteten ($1,96 S_w$) var 4,4 sekunder mellom prøverunden og første måling for begge testerne, og 3,5 sekunder for tester A og 3,6 sekunder for tester B, mellom første og andre måling. Dette samsvarer også med de absolutte gangverdiene som viste større forskjell i tid mellom prøverunden og første måling (1,6 og 1,7 sekunder for hver av testerne) enn mellom første og andre måling (0,2 sekunder for begge testerne) (tabell 2). Da ICC(1,1) kun var 0,81 og 0,84 for enkeltmålinger, beregnet vi reliabilitet av gjennomsnittet av to målinger. Ved utregning av gjennomsnitt av to målinger (mellom prøverunden og første måling, og mellom første og andre måling både for tester A og B), steg ICC(1,2) til 0,90 og 0,91 (tabell 5).

Figur 2 viser at en deltager brukte betydelig lenger tid enn de andre deltagerne på utførelse av testen. Da vi vet at en atypisk deltager kan innvirke på ICC-verdien, valgte vi å gjøre de samme beregningene av test-retest reliabilitet uten denne deltageren. ICC(1,1) sank da til 0,67 mellom prøverunden og første måling, og til 0,77 mellom første og andre måling. $1,96 S_w$ var ikke noe betydelig endret (tabell 6). Ved gjennomsnitt av to og to målinger steg ICC(1,2) til over 0,80.

Diskusjon

Oppsummert resultat

For eldre, hjemmeboende, skrøpelige personer testet av erfarne fysioterapeuter, viser TUG god relativ og absolutt reliabilitet både innen og mellom testerne og mellom to ulike observasjoner av samme person. God intratester- og intertesterreliabilitet betyr at et måleinstrument er godt egnet til å brukes av ulike testere. Test-retest reliabiliteten er også god, men viser noe større variasjon i gjennomsnittlig avvik, ICC og S_w verdier. Resultatet indikerer at tre repeterte målinger av TUG bør gjennomføres, fordi vi fant større målefeil mellom prøverunden og første måling enn mellom første og andre måling. Men først, hvilke metodiske forhold påvirker fortolkningen av resultatene? Og hvilken klinisk verdi har det studien frembringer?

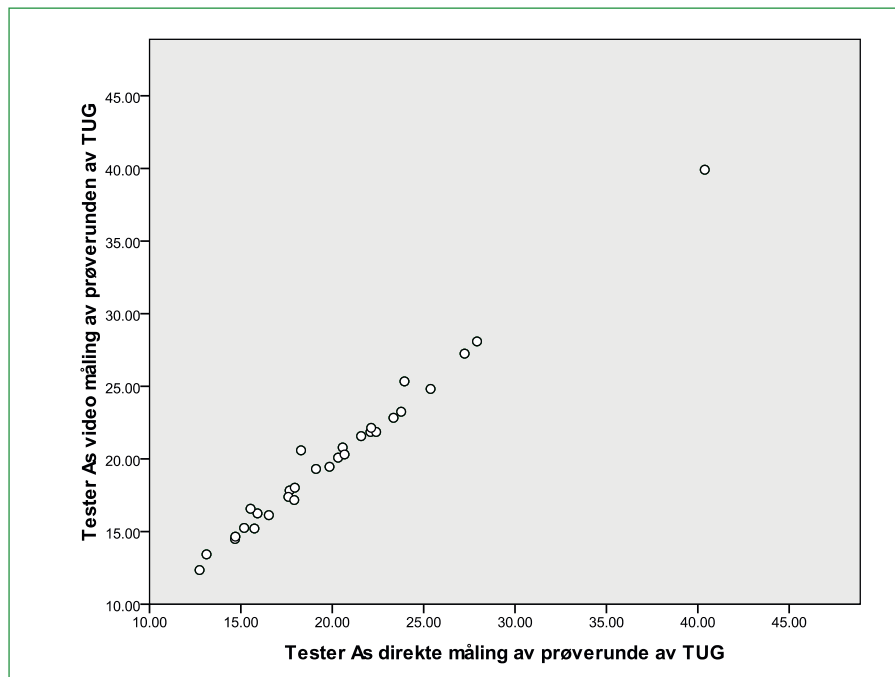
Metodediskusjon

Intern validitet: studiens styrker og begrensninger

To fysioterapeuter var involvert i å skåre testen, og en tredje fysioterapeut var ansvarlig for å gi all informasjon og instruksjon til deltagerne, samt ta video. Dette er ulikt testsituasjonen i klinikken, men ble valgt for å kunne teste intertesterreliabilitet, for å gi de to testerne samme oppgave under gjennomføringen, og for å sikre at alle deltagerne fikk lik informasjon og instruksjon. Resultatene på intertesterreliabilitet er derfor først og fremst uttrykk for reliabilitet i tidtaking mellom to testere. I klinikken hvor samme person vil gi instruksjon og ta tiden vil sannsynligvis intratester- og test-retest reliabiliteten bli enda høyere enn de allerede gode resultatene vi har i denne studien dersom det ble testet ut. Vi mener derfor at vår testprosedyre styrker studien.

Tester som innebærer bruk av tekst i instruksjon eller skåring bør oversettes fra originalspråket til det nye språket og så tilbakeoversettes og sammenlignes med den originale versjonen (27).

Etter studien gjennomførte vi en tilbakeoversettelse av instruksjonen benyttet i denne studien og sammenlignet den med den originale engelske versjonen av Podsiadlo og Richardson og med Sletvolds versjon, dette etter bestemte retningslinjer (27). TUG har vært benyttet i flere norske studier tidligere uten at denne oversettelsesprosedyren har vært gjennomført (11;12;16-18;20). Tilbakeoversettelsen viste at vi manglet instruksjon om ganghastighet fra Podsiadlo og Richardsons instruksjon og «uten assistanse» fra Sletvolds instruksjon. Det er grunn til å anta at en ytterligere presisering av ganghastighet, i enda en ny versjon av instruksjonen enn den vi har testet, vil forbedre test-retest reliabiliteten ytterligere. Dette da tidligere studier viser at selvvalgt tempo er et stabilt mål som gir god reliabilitet (18). Å inkludere ganghastighet i instruksjonen kan være en videre utbedring av protokollen for senere forskning.



FIGUR 2 Plottdiagram over absolutt verdiene fra tester As prøverunde måling direkte og fra video (målt i sekunder).

TABELL 2 Gjennomsnitt, standardavvik, minimum og maksimums skår for alle seks TUG målingene i absolutte verdier (sekunder). (N=30)

	Prøverunde				Prøverunde fra video				1. direkte måling			
	\bar{x}	SD	Min	Maks	\bar{x}	SD	Min	Maks	\bar{x}	SD	Min	Maks
Tester A	20,1	5,5	12,7	40,4	20,1	5,5	12,7	40,4	18,5	4,6	11,8	35,4
Tester B	20,2	5,5	12,7	40,6	19,7	5,5	12,0	39,5	18,5	4,6	11,5	35,6

	1. måling fra video				2. direkte måling				2. måling fra video			
	\bar{x}	SD	Min	Maks	\bar{x}	SD	Min	Maks	\bar{x}	SD	Min	Maks
Tester A	18,6	4,6	11,9	35,9	18,3	4,5	11,4	30,5	18,1	4,4	11,4	30,0
Tester B	18,4	4,6	12,0	35,5	18,3	4,4	11,5	30,2	18,2	4,3	11,5	30,0

\bar{x} = gjennomsnitt, SD = standard avvik, min = minimums verdien, Maks = maksimum verdien.

TABELL 3 Relativ og absolutt intratesterreliabilitet av TUG for testerne A og B (N=30): Relativ angitt med ICC(1,1) med 95 prosent konfidensintervall og absolutt med 1,96 Sw.

	ICC(1,1)	(95 % KI)	1,96 Sw
Tester As prøverunde direkte og fra video	0,99	0,99 – 0,99	±0,9
Tester As 1. måling direkte og fra video	0,99	0,99 – 0,99	±0,6
Tester As 2. måling direkte og fra video	0,99	0,99 – 0,99	±0,6
Tester Bs prøverunde direkte og fra video	0,99	0,98 – 0,99	±1,1
Tester Bs 1. måling direkte og fra video	0,99	0,99 – 0,99	±0,5
Tester Bs 2. måling direkte og fra video	0,99	0,99 – 0,99	±0,6

ICC(1,1) = Intraclass correlation coefficient (1,1), 95% KI = 95 % konfidensintervall, Sw = Within subject standard deviation.

Bohannon og andre studier viser at det varierer når tidtakingen begynner og slutter (3;11;16;19). Selve tidtakingen var i denne studien fra ordet «gå» til personen satt med sete ned på stolen igjen. Dette anbefaler vi også i protokollen. Start av tiden på instruksjonen «gå» mener vi er mer presis enn å vurdere når personen begynner å bevege seg. I tillegg mener vi at latenstiden fra «gå» til personen faktisk begynner å bevege seg er klinisk viktig. Selve testen ser ikke på enkeltoppgavene som inngår i testen (reise seg, gå, snu, gå eller snu og sette seg ned), men i klinisk praksis vil vi likevel anbefale at det observeres hvordan enkeltoppgavene utføres og eventuelt kommentere det i protokollen.

Ved tidtaking av de direkte målingene av TUG ble vanlige manuelle stoppeklokker brukt, mens ved skåringene fra video ble datamaskinen brukt som en manuell stoppeklokke. Her ser vi at det ville kunne styrket studien hvis en vanlig manuell stoppeklokke var brukt ved begge målingene. De gode resultatene på reliabilitet viser likevel at bruk av ulike stoppeklokker har hatt liten betydning.

Hvor mange ganger bør TUG gjennomføres for å få et mest mulig reliabelt mål? Bohannons studie viste at TUG har vært gjennomført med varierende antall repetisjoner siden testen ble etablert (15). I denne studien utførte deltagerne TUG tre ganger med tidtaking. Dette er ikke i tråd med verken Podsiadlo og Richardsons originale protokoll eller Sletvolds norske protokoll (3;20). En styrke ved vår studie er at vi har hatt muligheten til å undersøke reliabiliteten både mellom prøverunden og første måling, og mellom første og andre måling. Ut fra resultatene kan vi nå begrunne hvorfor TUG bør utføres tre ganger: en prøverunde og to ganger til med tidtaking gir minst målefeil. Vi fraviker dermed anbefalingen fra Podsiadlo og Richardsons og Sletvolds protokoller (3).

I test-retest reliabilitet er hensikten å se om deltagerens prestasjon er repeterbar

TABELL 4 Relativ og absolutt intertesterreliabilitet for tester A og B (N=30): Relativ angitt med ICC(1,1) med 95 prosent konfidensintervall og absolutt med 1,96 S_w.

	ICC(1,1)	(95 % KI)	1,96 S _w
Tester A og B prøverunde måling	0,99	0,99 - 1,00	±0,3
Tester A og B 1. måling	0,99	0,99 - 1,00	±0,3
Tester A og B 2. måling	0,99	0,99 - 1,00	±0,3

ICC(1,1) = Intraclass correlation coefficient (1,1), 95% KI = 95 % konfidensintervall, S_w = Within subject standard deviation.

TABELL 5 Relativ og absolutt test-retest reliabilitet for tester A og B (N=30): Relativ angitt med ICC(1,1) med 95 prosent konfidensintervall, og ICC(1,2) med 95 prosent konfidensintervall. Absolutt angitt med 1,96 S_w.

	ICC(1,1)	(95 % KI)	ICC(1,2)	(95 % KI)	1,96 S _w
Tester As prøverunde og 1. måling	0,81	0,65 - 0,91	0,90	0,79 - 0,95	±4,4
Tester As 1. og 2. måling	0,84	0,69 - 0,92	0,91	0,82 - 0,96	±3,5
Tester Bs prøverunde og 1. måling	0,81	0,63 - 0,91	0,90	0,78 - 0,95	±4,4
Tester Bs 1. og 2. måling	0,84	0,69 - 0,92	0,91	0,81 - 0,96	±3,6

ICC(1,1) = Intraclass correlation coefficient (1,1), 95% KI = 95 % konfidensintervall, ICC(1,2) = Intraclass correlation coefficient ICC(1,2), S_w = Within subject standard deviation.

TABELL 6 Relativ og absolutt test-retest reliabilitet for tester A og B: uten atypisk deltager (N=29). Relativ angitt med ICC(1,1) med 95 % konfidensintervall, og ICC(1,2) med 95 % konfidensintervall. Absolutt angitt med 1,96 S_w.

	ICC(1,1)	(95 % KI)	ICC(1,2)	(95 % KI)	1,96 S _w
Tester As prøverunde og 1. måling	0,67	0,42 - 0,83	0,81	0,59 - 0,91	±4,2
Tester As 1. og 2. måling	0,77	0,57 - 0,89	0,87	0,73 - 0,94	±3,4
Tester Bs prøverunde og 1. måling	0,67	0,41 - 0,83	0,80	0,58 - 0,91	±4,3
Tester Bs 1. og 2. måling	0,77	0,56 - 0,89	0,87	0,73 - 0,94	±3,4

ICC(1,1) = Intraclass correlation coefficient (1,1), 95% KI = 95 % konfidensintervall, ICC(1,2) = Intraclass correlation coefficient ICC(1,2), S_w = Within subject standard deviation.

(28). For å vurdere prestasjonene utførte vi tre repeterte målinger av testen. Det er ingen enighet blant forskere om tiden fra test til retest for denne type studier. Generelt anbefales det at tiden mellom testingene for test-retest reliabilitet bør vurderes på basis av deltagerens alder og helse og hvor anstrengende selve testen er. I tillegg bør en legge målingen til samme tidspunkt på døgn (28).

Utvalget i denne studien var relativt skrøpelige eldre mellom 75 og 92 år av begge kjønn. Skrøpelige eldre er i utgangspunktet mer utsatt for helsemessige endringer enn yngre, og har en mindre reserve. Dette indikerer at utvalg med relativt stor sannsynlighet ville variert mer i utførelsen av TUG enn det som var tilfelle dersom det hadde gått lang tid mellom test- og retestmålingene, både ut fra alder og helse i forhold til funksjon. En slik variabilitet er uttrykk for biologisk variabilitet og ikke reliabilitet av testen, og vi valgte derfor å ha kort tid mellom utførelse av test og retest. Studien kan derfor

ikke si noe om hvordan reliabilitet mellom dager er for denne gruppen eldre, men kun om TUGs korttidsrepeatabilitet.

Stolens høyde beskrevet i de forskjellige TUG studiene (3-6;9;12;15) varierer fra 43 til 48 cm. Podsiadlo og Richardson brukte en stol med armlener med setehøyde på 46 cm (3), som vi også benyttet i denne studien. Dersom en annen stolhøyde brukes, er det vesentlig at det presist rapporteres i protokollen, spesielt hvis TUG brukes for å måle endring i funksjon over tid. I den nye norske protokollen foreslår vi bruk av en stol med armlener som har en standardisert stolhøyde på 46 cm, som er en vanlig stolhøyde i Norge.

Ekstern validitet: bakgrunnsdata og utvalgsstørrelse

Ekstern validitet, eller generaliserbarhet, er avhengig av at studiens utvalg representerer et vidt spekter av funksjon innen den populasjonen testen er beregnet for (1). Det er en spredning i absoluttverdiene fra 11,4 til

40,6 sekunder som indikerer en variasjon i deltagerens gangfunksjon. Gangtiden spenner fra de som er selvstendige (under 20 sekunder) og de som avhengige av hjelp i flere dagligdagse aktiviteter (over 30 sekunder) (3). Aldersspredningen er fra 75 til 92 år og andel kvinner og menn tilsvarer andelen i aldersgruppen. Ut fra den funksjons- og alderspopulasjon som TUG er beregnet for, vurderer vi utvalget som rimelig representativt. Men da alle deltagerne ble rekruttert fra Aker universitetssykehus, er det ikke sikkert de representerer hele landet.

MMSE skår under 22 var et eksklusjonskriterium i denne studien. Dette var basert på erfaring med gjennomføring av intervensjonsstudie på samme populasjon (18) og som blant annet benyttet TUG til testing. Vi inkluderte ikke personer med større kognitiv svikt i studien, og resultatene kan derfor ikke generaliseres til denne gruppen.

Bruk av ICC og S_w

For intratester-, intertester- og test-retest reliabilitet har vi brukt to ulike statistiske mål: ICC for relativ reliabilitet, og S_w for absolutt reliabilitet.

ICC er den korrelasjonskoeffisienten som er beskrevet som den mest korrekte å bruke i forhold til relativ reliabilitet i reliabilitetsstudier (1;23). ICCs styrke er at de ser på både assosiasjon og enighet mellom målingene (23), er ikke influert av rekkefølgen i tallmaterialet (22) og er uavhengig av måleenheten.

Svakheter ved ICC er at disse verdiene kan bli affisert av variasjon i målingene på to måter. En lav ICC-verdi kan enten indikere at spredningen i verdiene mellom personene som testes er liten, eller at det er en stor spredning mellom målingene for den enkelte. En atypisk verdi kan gi en høyere ICC enn om den atypiske verdien ikke er tilstede, men kan også gi en lavere ICC avhengig av om variabiliteten mellom målingene eller individer er stor eller liten. Test-retest reliabiliteten i denne studien viste dette fenomenet. Én deltager brukte betydelig lenger tid enn de 29 andre deltagerne, og påvirker ICC-verdiene i positiv retning. Ved fjerning av deltageren fra datasettet så vi lavere ICC-verdier (tabell 5 og 6).

S_w var valgt som det primære målet for absolutt reliabilitet da den gir uttrykk for målefeilen til de repeterte målingene for hvert individ, og da S_w oppgis i den aktuelle måleenheten (24;25). En begrensing ved S_w

er at den ikke kan sammenlignes mellom mål med ulike måleenheter, men dette var ikke aktuelt her.

Tidlige reliabilitetsstudier varierer med hensyn til reliabilitetsmål som er benyttet (3), og inntil de siste årene er det ofte kun mål for relativ reliabilitet som er rapportert (6;19;29). Valg av type ICC i forhold til hvilken type reliabilitet som undersøkes varierer også (6), og valgene er ikke alltid forenlig med Shrout og Fleiss sine anbefalinger (23). Noen artikler oppgir kun gjennomsnittet av flere målinger istedenfor rapportering av enkeltmålinger (19;29). Gjennomsnittet av flere målinger gir ofte en høyere ICC-verdi enn en enkelt måling, da gjennomsnittet fjerner tilfeldighetene i målingene (30), som kan være grunnen for å oppgi gjennomsnittet av flere målinger. I noen nyere studier er det mer konkret rapportert hvilken ICC som er benyttet, og absolutt reliabilitet er også rapportert (30).

Resultatdiskusjon

TUG viser i denne studien god intratester-, intertester- og test-retest reliabilitet. At begge testerne var erfarne fysioterapeuter og hadde brukt TUG i praksis kan være en grunn til de gode resultatene. Det er usikkert om reliabiliteten hadde vært så god dersom en uerfaren tester hadde vært inkludert. Vi kan ikke si noe om dette ut fra våre resultater, men det fremhever at fysioterapeutene bør gjøre seg kjent med testen før den benyttes i klinisk praksis.

Test-retest reliabilitet har fokus på å vurdere biologisk variabilitet (1), og om testen er stabil over tid. Ikke uventet er den relative reliabiliteten av test-retest noe dårligere enn intratester- og intertesterreliabilitet, og målefeilen noe større. Denne variabiliteten indikerer en variasjon i deltagerens utførelse. ICC-verdiene er større ved gjennomsnittet av to målinger enn ved én måling. Det siste har vært benyttet i andre studier tidligere (6;18). Vår studie viser en liten forskjell (1,1 sekunder) i målefeil mellom prøverunden og første måling, og mellom første og andre måling (tabell 5). En så liten forskjell kan virke ubetydelig, men i klinisk sammenheng kan det være vesentlig for å si noe om en persons endring av funksjon. Når TUG brukes for å måle endring mellom to forskjellige måletidspunkt, må endringen ut fra våre data være på over 3,6 sekunder for å kunne si med 95 prosent sannsynlighet at den er reell.

Betydning for praksis

En test som skal benyttes i klinisk praksis må være reliabel innen og mellom testere og for de samme personene over tid for det geografiske og språklige området testen skal benyttes i. For at en test i tillegg skal være nyttig må endringene som sees fra test til retest være forårsaket av egenskaper hos personen, og ikke skyldes testeren eller testen.

Den nye norske protokollen er i denne studien presisert og reliabilitetstestet på norsk og med et utvalg med en variasjon i funksjon, alder og kjønn vi mener representerer den populasjonen med balanse- og gangproblemer som TUG er beregnet for. Dette gjør at testen er god og kan benyttes i klinisk praksis og forskning.

Konklusjon

Studien av den nye norske versjonen av TUG, hvor erfarne fysioterapeuter tester eldre personer med fall- og gangproblemer, viser god intratester-, intertester- og test-retest reliabilitet. Resultatene viser en liten variasjon i test-retest tidene som indikerer varierende prestasjon innad hos hver deltager fra en måling til det neste. Resultatene viser en høyere relativ reliabilitet ved benyttelse av gjennomsnittet av andre og tredje måling. Derfor anbefaler vi i den nye norske protokollen (vedlegg 1) at testen gjennomføres tre ganger med et par minutters mellomrom: en ikke gjeldende prøverunde og to ganger med tidtaking hvor gjennomsnittet gjøres gjeldende. Den absolutte reliabiliteten viser at når TUG brukes for å måle endring mellom to forskjellige måletidspunkt, må endringen være på over 3,6 sekunder for å være trygg på at den er reell.

Vi mener som andre, at TUG er en enkel test som gir et reliabelt bilde av skjøre eldre personers balanse, ganghastighet og generelle funksjon, og at den kan benyttes i alle kliniske settinger, inkludert hjemme hos brukeren (3). Vi foreslår en mer presis instruksjon enn tidligere i den nye protokollen for TUG (vedlegg 1). Den nye protokollen gjør at norske fysioterapeuter mer reliabelt enn tidligere kan bruke TUG som måleverktøy.

Viktig videre forskning på den nye norske protokollen vil være å vurdere test-retest reliabilitet over en lengre periode, med instruksjon i forhold til ganghastighet på en populasjon med dårligere kognitiv funksjon enn her og om den eventuelt kan predikere fall i den norske populasjonen.

Finansiering

Studien har fått delvis økonomisk støtte fra Høgskolen i Oslo (HiO), Praksismidler mellom HiO og Aker universitetssykehus. Forøvrig er prosjektet finansiert av HiO og Aker universitetssykehus i form av tid.

Takk

En stor takk til alle pasientene som deltok i studien. Uten dem ville ikke denne studien vært mulig. En stor takk også til fysioterapeutene Charlotta Hamre og Trine Strøm for deltagelse i pasientrekruttering, registrering og for å være med i alle testsituasjonene. I tillegg en stor takk til fysioterapistaben og de andre faggruppene ved Aker universitetssykehus, avdeling for rehabilitering og geriatri, for meget god hjelp med rekruttering, sikring av deltagerne, bruk av lokaler og godt samarbeid.

Litteratur

1. Domholdt. Physical therapy research : principles and applications. Philadelphia: Saunders; 2000.
2. Finch, Brooks, Statford, et al. Physical rehabilitation outcome measures: a guide to enhanced clinical decision-making. Baltimore, MD: Lippincott Williams & Wilkins; 2002.
3. Podsiadlo D, Richardson S. The Timed Up and Go - A Test of Basic Functional Mobility for Frail Elderly Persons. Journal of the American Geriatrics Society 1991 39 (2): 142-8.
4. Rockwood K, Awalt E, Carver D, et al. Feasibility and measurement properties of the functional reach and the timed up and go tests in the Canadian study of health and aging. Journal of Gerontology: Medical Sciences 2000; 55 (2): 70-3.
5. Nordin E, Rosendahl E, Lundin-Olsson L. Timed "Up & Go" test: reliability in older people dependent in activities of daily living--focus on cognitive state. Physical Therapy 2006; 86 (5): 646-55.
6. Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. Physical Therapy 2002; 82 (2): 128-37.
7. Botolfson P, Helbostad JL, Moe-Nilssen R, et al. Reliability and concurrent validity of the Expanded Timed Up-and-Go test in older people with impaired mobility. Physiotherapy Research International 2008; 13 (2): 94-106.
8. Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the "get-up and go" test. Archives of Physical Medicine and Rehabilitation 1986; 67 (6): 387-9.
9. Schoppen T, Boonstra A, Groothoff JW, et al. The timed "up and go" test: Reliability and validity in persons with unilateral lower limb amputation. Archives of Physical Medicine and Rehabilitation 1999; 80 (7): 825-8.
10. Berg KO, Wooddauphinee SL, Williams JL. Measuring Balance in the Elderly - Validation of An Instrument. Canadian Journal of Public Health-Revue Canadienne de Sante Publique 1992; 83: 7-11.
11. Bergland A, Jarnlo GB, Wyller TB. [Self-reported walking, balance testing and risk of fall among the elderly]. Tidsskrift for Den Norske Legeforening 2006 12; 126 (2): 176-8.
12. Langhammer B, Lindmark B. Performance-Related Values for Gait Velocity, Timed Up-and-Go and Functional Reach in Healthy Older people and Institutionalized Geriatric Patients. Physical and occupational therapy in Geriatrics 2007; 25 (3): 55-69.
13. Freter SH, Fruchter N. Relationship between timed 'up and go' and gait time in an elderly orthopaedic rehabilitation population. Clinical Rehabilitation 2000; 14 (1): 96-101.

Title: Reliability of the Norwegian version of Timed Up and Go (TUG)

Abstract

- **Aim:** To test the intrarater-, interrater- and test-retest reliability of the Norwegian protocol of the «Timed Up and Go» (TUG).
- **Design:** Cross-sectional method study.
- **Material and Method:** Thirty persons (20 women and 10 men) 75 years or older (mean 82.5) completed TUG three times. To physiotherapists scored the participants. Intra-class correlation coefficient (ICC) was computed to express measurement variability relative to total variability, and within-subject standard deviation (S_w) calculated as an expression of absolute variability.
- **Result:** The mean times were from 18.1 to 20.2 seconds. Intrarater- and interrater reliability expressed as ICC(1,1) showed 0,99 for all the measurements. $1,96 S_w$ was from $\pm 0,53$ to $\pm 0,57$ seconds for intrarater reliability and $\pm 0,3$ seconds for interrater reliability. All ICC(1,1) values for the test-retest reliability was higher than 0,81. $1,96 S_w$ was between $\pm 3,5$ to $\pm 4,4$ seconds. When calculating the mean of two measurement, ICC(1,2) the reliability increased to over 0,90.
- **Conclusion:** This study demonstrates that TUG, on fragile elderly persons performed by experienced therapists, has high intrarater-, interrater- and test-retest reliability. We recommend that the mean of measurement two and three of the new Norwegian TUG version is used in the clinic.
- **Keywords:** outcome measure, clinical protocol, reliability.

14. Hughes C, Osman C, Woods A. Relationship among performance on stair ambulation, Functional Reach, and Timed Up and Go tests in older adults. Issues on Aging 1998; 21: 18-22.
15. Bohannon RW. Reference values for the timed up and go test: a descriptive meta-analysis. Journal of Geriatric Physical Therapy 2006; 29 (2): 64-8.
16. Bergland A, Jarnlo GB, Laake K. Predictors of falls in the elderly by location. Aging Clinical and Experimental Research 2003; 15 (1): 43-50.
17. Thrane G, Joakimsen RM, Thornquist E. The association between timed up and go test and history of falls: the Tromsø study. British Medical Journal Geriatrics 2007; 7: 1-7.
18. Helbostad JL, Sletvold O, Moe-Nilssen R. Effects of home exercises and group training on functional abilities in home-dwelling older persons with mobility and balance problems. A randomized study. Aging Clinical and Experimental Research 2004; 16 (2): 113-21.
19. Shumway-Cook A, Brauer S, Woolacott M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. Physical Therapy 2000; 80 (9): 896-903.
20. Sletvold, Engedal, Tilvis, et al. Geriatrisk utredning i Norden: nordiske retningslinjer for spesialhelsetjenesten i geriatri. Oslo: Den norske legeforening; 1997; 117(24): 3540-1.
21. Moe-Nilssen R, Nordin E, Lundin-Olsson L. Criteria for evaluation of measurement properties of clinical balance measures for use in fall prevention studies. Journal of Evaluation in Clinical Practice 2008; 14 (2): 236-40.
22. Bland JM, Altman DG. Statistics Notes: Measurement error and correlation coefficients. British Medical Journal 1996; 313 (7048): 41-2.
23. Shrout PE, Fleiss JL. Interclass Correlations: Uses in Assessing Rater Reliability. Psychological Bulletin 1979; 86 (2): 420-8.
24. Bland JM, Altman DG. Statistics Notes: Measurement error. British Medical Journal 1996; 313 (7059): 744.
25. Bland JM, Altman DG. Statistics Notes: Measurement error proportional to the mean. British Medical Journal 1996; 313(7049): 106.
26. Munro. Statistical methods for health care research. Philadelphia: Lippincott; 2001.

27. Streiner, Norman. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press; 2008.
28. Deitz JC. Reliability. Physical & Occupational therapy in Pediatrics 1989; 9(1): 125-69.
29. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. Clinical Rehabilitation 1998; 12 (3): 187-99.
30. Helbostad JL, Askim T, Moe-Nilssen R. Short-term repeatability of body sway during quiet standing in people with hemiparesis and in frail older adults. Archives of Physical Medicine and Rehabilitation 2004; 85 (6): 993-9.

Vedlegg 1 - neste side

VEDLEGG 1

TUG

THE TIMED «UP & GO» TEST

D Podsiadlo, S Richardson. J Am Geriatr Soc 1991;39:142-148.

Bearbeidet til norsk av professor dr. med. Olav Sletvold.

Videre bearbeidet og reliabilitets testet av Olav Sletvold, Jorunn L. Helbostad, Charlotta Hamre og Pernille Botolfsen. Publisert i Fysioterapeuten 5/2010.

Navn: _____ F. nr: _____

Fødselsår/dato: _____ Dato utfylt: _____

Utfylt av: _____

Bruksområde

Screeningtest for bevegelseshemmede. Informativ vedrørende personens balanse, ganghastighet og generell funksjonsevne.

Forberedelser

Pasienten har på gode innesko, og kan bruke det ganghjelpemiddel hun/han vanligvis bruker. Hun/han sitter tilbaketil i stolen mens armene hviler på armlenene.

Utstyr

Stol med armlener, setehøyde 46 cm. Stoppeklokke. Merke med 50 cm farget tape på gulvet 3 meter fra fremre stolben.

Instruksjon

«Etter ”klar gå” så reiser du deg opp, går forbi den røde streken (den aktuelle fargen på tapen på gulvet), snur, går tilbake til stolen og setter deg ned igjen.» Testen gjennomføres tre ganger. Først en prøverunde og så to ganger med tidsmåling og det beregnes gjennomsnitt av de to målingene. Instruksjonen på 1. og 2. måling er kun ”klar-gå”. Instruktøren sikrer at pasienten ikke faller.

Tidtaking

Start av tiden på instruksjonen ”gå” til personen sitter med sete ned på stolen igjen.

1. måling: _____ 2. måling: _____ Gjennomsnitt: _____

Kommentarer til utførelsen: _____

Vurdering

< 20 sekunder: Trenger ikke personhjelp ved forflytning.

> 30 sekunder: Trenger ofte personhjelp i forflytning. Kan som regel ikke gå i trapper og gå ute alene. Husk, pasienten må oppfatte og huske instruksjonen!