Sammendrag

Bakgrunn

Temperatur kan måles på mange ulike måter. Ved triage i akuttmottak benyttes enkle, non-invasive termometre med ulik validitet og reliabilitet. Hensikten med denne studien var å finne ut om temporaltermometeret avviker fra den rektale gullstandard, og om det er et godt klinisk instrument. Vi har sett på validitet gjennom å beregne sensitivitet og spesifisitet, samt beregnet reliabiliteten; apparatets evne til å reprodusere tilsvarende målinger hos samme individ.

Material og metode

Temperatur målt temporalt ble sammenlignet med rektaltemperatur hos 70 tilfeldig utvalgte pasienter i akuttmottaket ved St.Olavs Hospital. Det ble gjort to målinger med hvert apparat hvilket gav grunnlag for vurdering av reliabilitet og validitet.

Resultater

Temporalmålet gav i gjennomsnitt -0,49°C lavere temperatur enn rektalmålet (95% KI -0,62, -0,36), p<0,01). Sensitiviteten for å oppdage feber ved temporalmåling var 47,6%. Negativ prediktiv verdi (NPV) var 81,7% og positiv prediktiv verdi(PPV) var 100%. Hos de febrile pasientene hadde temporalmålet et gjennomsnittlig avvik på -0,87°C sammenlignet med temperaturen målt rektalt. Gjennomsnittlig avvik hos den afebrile gruppen(n=49) var -0,33°C. En kalibrering av temporalmålet hvor en legger til 0,5°C på alle temporalmålinger, ville gitt en sensitivitet på 57,1%, en PPV på 70,6% og NPV på 83%. Gjennomsnittlig avvik mellom første og andre temporalmåling var -0,056°C (95% KI: -0,1, -0,01, p=0,02) og gjennomsnittlig avvik mellom første og andre rektalmåling var 0,023°C (95% KI: 0,01, 0,06, p<0,18).

Tolkning

Studien viste at temporalmålet har et avvik fra den rektale gullstandard som er innenfor det en kan anse som akseptabelt for klinisk bruk. Apparatet har en høy reliabilitet og er svært praktisk i bruk. Det er samtidig viktig å kjenne til apparatets begrensninger. Hos febrile pasienter viser vår studie at avviket fra gullstandarden er større enn hos de afebrile pasientene, hvilket kan resultere i feildiagnostikk og feilbehandling av et stort antall pasienter. Hadde man kalibrert temporalmålet opp 0,5°C ville en ha fått bedre overenstemmelse med rektalmålet.

Introduksjon

Når pasienter kommer inn i et akuttmottak, er det essensielt å få oversikt over de vitale parametere, som puls, respirasjonsfrekvens, blodtrykk og temperatur. Å overse feber kan være kritisk for den enkelte pasient[1]og ved vurdering av infeksjonspasienter er temperaturmåling sentralt. Et feilaktig underestimat av temperatur kan resultere i at sepsis ikke oppdages, antibiotikakurer startes for sent og blodkulturer ikke tas.[2] For å kunne tilby adekvat utredning og behandling er det nødvendig med pålitelige måleinstrumenter, og det varierer fra sykehus til sykehus hvilket som benyttes til temperaturmåling.

Det rektale kvikksølvtermometeret har lenge vært det viktigste instrumentet for å måle kroppstemperatur. Det har vært en viktig referanse i klinisk diagnostikk, men er ikke lenger i salg, både av miljømessige og funksjonelle årsaker. I dagens akuttmottak er det nå erstattet med non-invasive infrarøde og digitale instrumenter. Disse kan enkelt og raskt måle pasientens temperatur oralt, i pannen(temporal arterie temperatur - TAT), axillært og i øret. [3] Utfordringen er hvor pålitelig disse instrumentene er. [4]

Furlong et al. presenterte i 2014 en studie hvor de konkluderte med at temperatur målt i a.pulmonalis(PAT) er mest lik kroppens kjernetemperatur, fordi blandet venøst blod gir det beste bildet av hypothalamus´ termoregulering. [5] Denne avanserte metoden er åpenbart for invasiv utenfor en intensivavdeling. Andre alternative invasive målemetoder er øsofagal-, blære og rektalmåling. I klinisk praksis er rektaltemperatur mest benyttet, og er ansett som den minst invasive av de nevnte målemetoder. En litteraturstudie gjennomført av Fulbrook (1993) demonstrerte at også denne metoden er tilnærmet lik kroppens kjernetemperatur.[6]

Palpasjon av pannen har lenge vært brukt som pekepinn for å angi feber da a.temporalis superficialis ligger her. A. temporalis superficialis forløper foran øret, opp i temporalområdet og ender i en gaffelformasjon mellom huden og kraniet, ved siden av og like under øyebrynet. Denne arterien har en relativt kort avstand til hjertet, og har en høy og tilnærmet konstant blodflow. Tanken bak TAT-måling er at den skal gi et bilde av temperaturen i hjertet, og dermed være tilnærmet lik en PAT-måling. [6, 7]

Flere studier er gjort for å teste de ikke-invasive instrumentene opp i mot rektaltemperatur. [8, 9] En rapport fra Kunnskapssenteret (2009) avdekket imidlertid at det ikke er gjort studier på den diagnostiske nøyaktighet til pannemåling, såkalt temporal arterie temperatur(TAT), i forhold til rektalmåling. [10] Lokalt ved St.Olavs Hospital ble det rapportert om stor diskrepans mellom målt TAT og rektaltemperatur. Hvor stor denne diskrepansen er, er ikke kjent. Dette gav grunnlaget for vår studie.

Hensikten med denne studien var å finne ut om temporaltermometeret avviker fra den rektale gullstandard, og om det er et godt klinisk instrument. Vi har sett på validitet gjennom å beregne sensitivitet og spesifisitet, samt beregnet reliabiliteten; apparatets evne til å reprodusere tilsvarende målinger hos samme individ.

Metode

Studiepopulasjon
Studien inkluderte voksne pasienter til vurdering i akuttmottaket ved St. Olavs Hospital uavhengig av kontaktårsak, kjønn og alder. Det ble gjort målinger av totalt 70 pasienter i tidsperioden september 2015 til februar 2017. Gjennomsnittsalderen for pasientene var 66 år (19-100 år) og median 71 år. Av forsøkspersonene var det 28 kvinner og 42 menn. Pasienter der deltagelse i studien ville forsinket tidskritisk utredning og behandling, samt pasienter som ikke samarbeidet eller av andre årsaker motsatte seg deltakelse, ble ekskludert fra studien.

Datainnsamling
To faste medisinerstudenter har fulgt triage-sykepleiere i akuttmottaket. TAT-målingene er gjennomført av studentene. For å gjenspeile gjeldende praksis har pasientene selv fått velge om den rektale temperaturmålingen gjøres av studentene eller pasienten selv. Det er ikke tatt hensyn til dette i arbeidet med tallmaterialet. Brukerveiledninger fra produsentene er nøye studert, og den praktiske gjennomføringen av målingene har fulgt disse for å unngå systematiske feil i studien. Forsøket er godkjent av regional etisk komité.

Det temporale termometeret var et TemporalScanner modell TAT-5000(Exergen) infrarødt termometer. Rektaltermometeret var av typen Terumo ET-C405H. Etter gjeldende praksis ble temporal termometer tilhørende hvert enkelt mottaksrom benyttet, og det var totalt 10 rom. 4 ulike rektalmålere ble benyttet. På hver pasient ble det utført to TAT- og to rektalmålinger(R1 og R2). Disse ble gjennomført over et tidsintervall på 2 minutter.

Dataanalyse
Temporalmålets validitet ble beregnet ut i fra første måling(TAT1 og R1) for hver målemetode. Rektaltemperaturen ble regnet som gullstandard. Avvik her ble beregnet som differansen mellom TAT og R1, slik at en negativ verdi innebærer at temporaltemperaturen er lavere enn den rektale målingen.

Reliabiliteten til de to apparatene ble beregnet ut i fra to repeterte målinger hos samme individ (TAT1-TAT2 og R1-R2). Programmene Excel (Ver. 14.4.6 2011) og SPSS (Ver.24) ble benyttet for analyse.

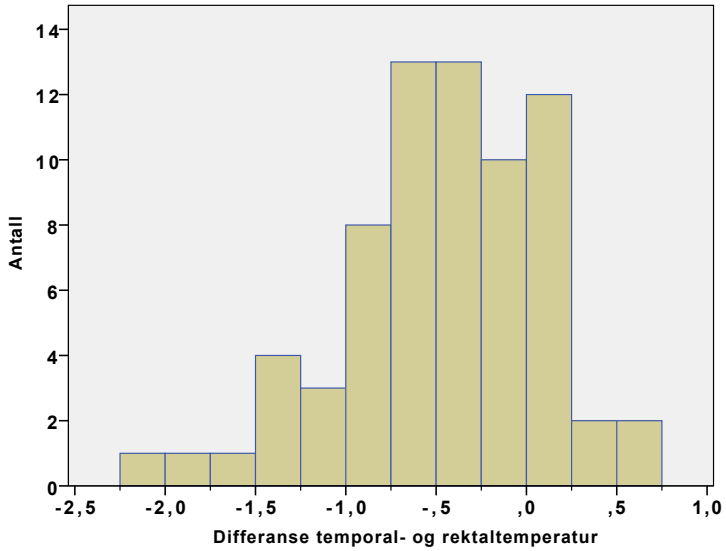
Vi ønsket å se om det forelå en klinisk og statistisk signifikant forskjell mellom TAT-måling og den rektale gullstandard, herunder analysere hvorvidt det var forskjell mellom febrile og afebrile pasienter. Feber ble definert som rektaltemperatur ≥38,0°C.[11] Student T-test ble brukt for å regne ut gjennomsnittsverdier.[12] Differansen mellom resultatene (TAT1 og R1) ble presentert som gjennomsnittsverdier med 95% konfidensintervall(KI). Grafisk fremstilling er gjort i samsvar med Bland&Altman for sammenligning av to diagnostiske tester, og scatterplot m/likhetslinje er benyttet for presentasjon av reliabilitet. [13] En forskjell lik eller større enn 0,5 grad ble ansett som klinisk signifikant. I forkant av studien er det i samråd med eksperter definert akseptable verdier for bias og presisjon på ±0,5°C ved bruk av et non-invasivt instrument som substitutt for invasiv måling.

Studiens utvalgsstørrelse ble beregnet i samråd med statistiker. Styrke og signifikansnivå ble satt til 0,05. Gjennomsnittlig differanse mellom TAT1 og R1 ble anslått til 0,5°C. Standardavvik (SD) ble anslått, sannsynligvis noe høyt, til 1°C for å unngå å underestimere spredning, og på den måte inkludere for få i studien. Kalkulator for utvalgsstørrelse gav med disse rammene et anslag på 58 pasienter, under betingelse om normalfordeling. Vi økte med 20% for mulig skjevfordeling i resultatene. Utvalgsstørrelse(n) ble da 70 pasienter.

Resultater

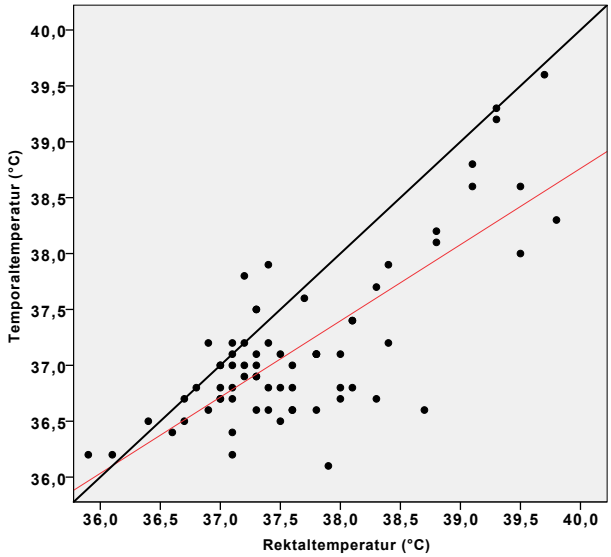
Validitet

Til sammen ble det gjort TAT- og rektalmålinger på 70 pasienter. Temporalmåleren målte i gjennomsnitt 0,49°C lavere temperatur enn rektaltermometeret (95% KI -0,62,-0,36, p<0,01). (fig.1)



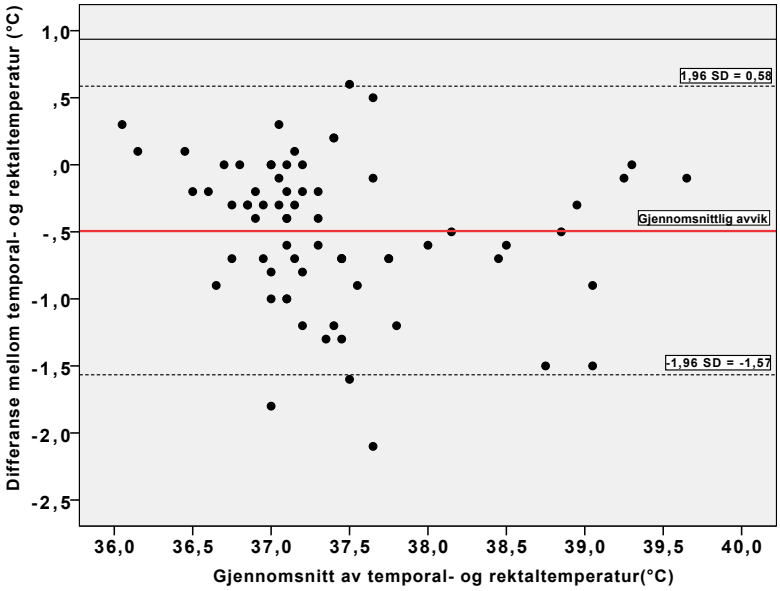
Figur 1: Histogram over fordeling i differansen mellom temporal- og rektaltemperatur (x-aksen) og antallet målinger med gitt temperaturavvik (y-aksen).

Hos de febrile pasientene (≥38°C) hadde temporalmålet et gjennomsnittlig avvik på -0,87°C (95% KI: -1,12, -0,62, p<0,001) sammenlignet med temperaturen målt rektalt. Gjennomsnittlig avvik hos den afebrile gruppen (n=49) var -0,33° (95% KI: -0,47, -0,20) (p<0,001). Scatterplottet viser grafisk hvordan nesten alle punktene ligger under likhetslinjen (x=y), og gir et bilde av at temporalmåleren gjennomgående måler lavere temperatur enn rektaltermometeret. Regresjonslinjen har en positiv stigningskoeffisient som er lavere enn én, og dette tyder på at avviket mellom de to målemetodene blir progressivt større ved en gradvis økning av temperatur målt rektalt. (fig.2)



Figur 2: Scatterplottet viser korrelasjonen mellom rektaltemperatur (x-aksen) og temporaltemperatur (y-aksen). Rød linje er regresjonslinje og svart linje gir x=y.

Dette Bland Altman-plottet viser forskjellen mellom de to målemetodene, og gir oss en grafisk fremstilling på hvordan differansen er distribuert i datasettet. (fig.3)



Figur 3: Sammenhengen mellom gjennomsnittet av de to målemetodene (x-aksen) og differansen mellom temporal- og rektaltemperatur (y-aksen), presentert i et Bland & Altman-plot.

Sensitivitet og spesifisitet

21 av 70 (30%) pasienter hadde feber (≥38,0°C) ved rektalmåling, og 10 av 70 ved TAT-måling (tab.1). Sensitiviteten for å oppdage feber med TAT-måling var 47,6% (95% KI 35,9, 59,6%) og spesifisiteten 100%. PPV var 100% og NPV var 81,7% (95% KI 72,6, 90,7%). Kappa-koeffisient var noe lav på 0,56, sannsynligvis som følge av få febrile pasienter i utvalget. McNemar-tall 0,001 viser at det forelå systematisk forskjell.

Rektalmålinger

		< 38°C	≥38°C	Totalt
Temporalmålinger	< 38°C	49 (70%)	10 (14,3%)	59 (84,3%)
	≥ 38°C	0 (0%)	11 (15,7%)	11 (15,7%)
	Totalt	49 (70%)	21 (70,0%)	*70 (100%)*

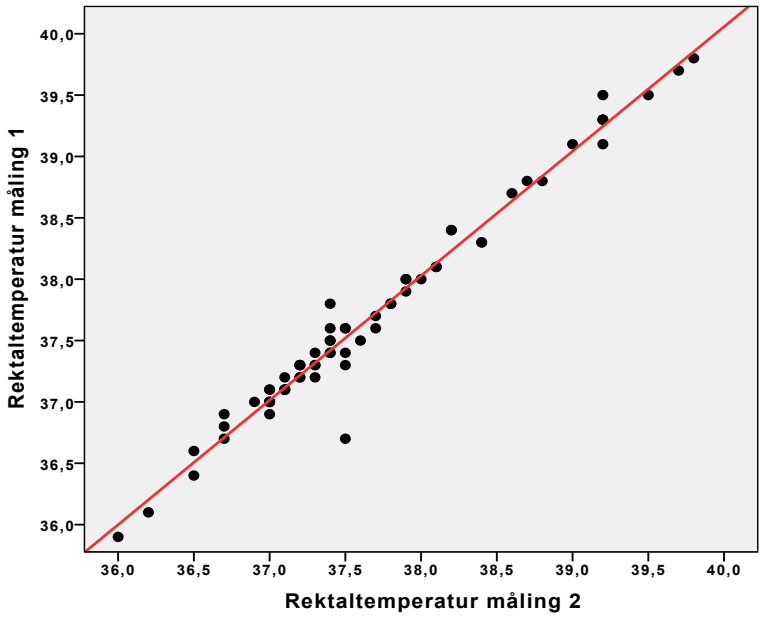
Tabell 1: Fordeling av pasientpopulasjonen i afebrile og febrile grupper.

Kalibrering av TAT-måleren

Dersom vi legger til 0,5°C på TAT-målingene hadde 17 av 70 pasienter feber ved TAT-måling. Sensitiviteten for å oppdage feber ble da 57,1% (95% KI 45,5, 68,7%) og spesifisiteten 89,8% (95% KI 82,7, 96,9%). PPV ble 70,6% (95% KI 59,9, 81,2%) og NPV 83,0% (95% KI 74,2, 91,8%).

Rektalmålerens reliabilitet

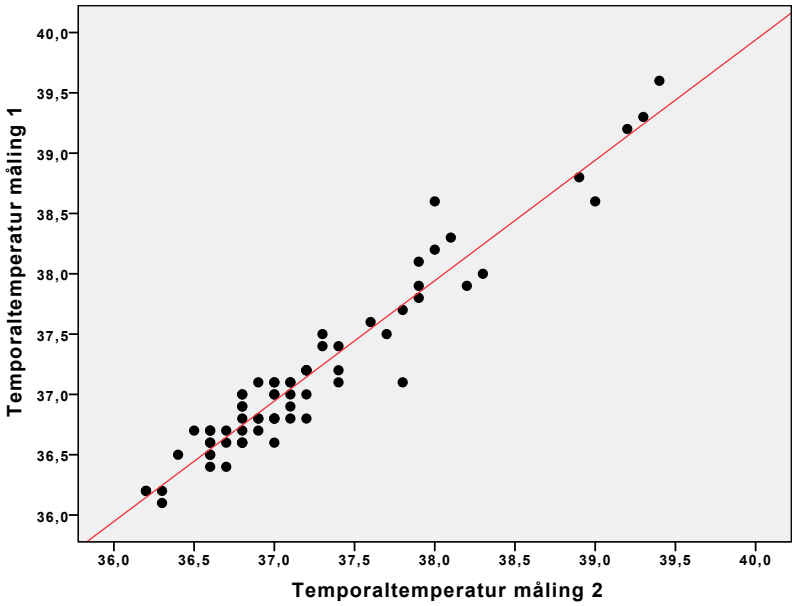
Gjennomsnittlig avvik mellom første og andre rektalmåling var 0,023 (95% CI: 0,01, 0,06, p<0,18). (fig.4)



Figur 4: Her fremstilles et plot som viser rektalmålets reliabilitet.Y er tilnærmet lik X i hele datasettet.

Temporalmålerens reliabilitet

Gjennomsnittlig avvik mellom første og andre var her -0,056 grader (95% CI: -0,1, -0,01, p=0,02). (fig.5)



Figur 5: Plottet over fremstiller temporalmålerens reliabilitet. En ser at det er noe større avvik mellom Y og X her enn hos rektalmåleren.

Diskusjon

Studien viser at infrarød temperaturmåling over temporalarterien gir et lavere temperaturestimat enn rektale temperaturmål hos pasienter i akuttmottak. For hele utvalget var det et statistisk signifikant gjennomsnittlig avvik på -0,49°C, altså under den forhåndsbestemte grense for klinisk signifikant avvik (±0,5°C). Dette skulle tilsi at apparatet er tilstrekkelig for klinisk praksis. Hvis man ser på sensitiviteten for TAT-målet var denne på 47,6%, og det påviste altså feber hos kun omlag halvparten av de febrile pasientene. PPV og NPV var henholdsvis 100% og 81,7%, og av alle pasientene som var afebrile ved temporalmål, var dermed omlag en femtedel av disse febrile etter den rektale gullstandard. Avviket for hele utvalget er altså *ikke* klinisk signifikant.

Ved å dele populasjonen i en febril og afebril gruppe basert på første temperatur målt rektalt, fant vi at avviket var klinisk signifikant hos den febrile gruppen (-0,87°C), og noe mindre og ikke klinisk signifikant (-0,33°C) hos den afebrile gruppen. Bland Altman-plottet (fig. 3) og scatterplot med regresjonslinje (fig. 2) gir et bilde av at forskjellene mellom målemetodene er større ved høyere temperatur, og lavere ved lavere temperatur. Dette taler for at temporalmålet er et godt diagnostisk verktøy for afebrile pasienter, men ikke tilstrekkelig for de febrile. Hos de aller laveste temperaturene er den målte temporale temperatur høyere enn den rektale. Hvordan dette faktisk er for pasienter med lavere temperaturer svarer ikke denne studien på, da studiens kliniske temperaturintervall ikke omfatter hypoterme pasienter.

Infrarøde temporalmålere er også i tidligere studier vist å underestimere temperaturer sammenlignet med temperatur målt i rektum. Dybwik, Nielsen og medarbeidere fant at infrarød temporalmåling i gjennomsnitt lå 0,3°C under temperaturen målt i rektum hos voksne intensivpasienter, og hadde en sensitivitet for å oppdage feber på 53%. [14] Resultatene i vår studie underbygger funn gjort i denne studien.

En kan tenke at en kalibrering av temporalmålet ville være en god løsning. Ved å legge til 0,5°C for å kompensere for det gjennomsnittlige avviket, ville dette gitt en sensitivitet på 57,1%. Dette går på bekostning av spesifisiteten, som blir 89,9%. Man ville med andre ord fanget opp omlag 10% flere av de febrile, men samtidig også feilaktig ansett 10% av de afebrile som febrile. NPV ville blitt marginalt øket til 83%, og PPV redusert til 70,3%. I vår studie var fordelingen i populasjonen tilfeldig utvalgt, der den febrile og afebrile gruppen utgjorde henholdsvis 30% og 70%, og det er ikke urimelig å anse dette som normalen i et akuttmottak. Resultatet av en slik kalibrering vil derfor være at det totale antallet man *feilaktig* inkluderer i febergruppen er relativt større enn de man *rettmessig* fanger opp som følge av kalibreringen.

Som et ledd i studien ble det gjennomført en reliabilitetstest av de to måleapparatene ved å se på deres evne til å reprodusere tilsvarende målinger hos samme individ. Gjennomsnittlig avvik blant rektalmålene var 0,023°C, og dette avviket var ikke statistisk signifikant. Temporalmålerne hadde et statistisk signifikant gjennomsnittlig avvik på 0,056°C. Dette avviket har sannsynligvis ikke klinisk betydning, men det viser at temporalmålet har lavere reliabilitet enn rektaltermometeret.

Temporalmålet er et lett anvendelig apparat. Det krever lite opplæring, er svært praktisk i bruk og medfører marginalt med stress og ubehag for pasienten. Apparatet er også lett å vedlikeholde. Disse momentene tatt i betraktning, må man ikke la apparatets praktiske egenskaper gå ut over pasientsikkerheten. Validiteten til apparatet er åpenbart det viktigste, og er i denne studien funnet å være tvilsom.

De gjennomsnittlige avvikene fra gullstandard er betydelige, og sensitiviteten for å oppdage feber så lav at dette vil kunne resultere i feildiagnostikk hos enkelte pasienter, og det er i den febrile gruppen vi finner de største avvikene. Lav sensitivitet og et avvik nært opp mot hva anses som klinisk signifikant, er funn som må vektlegges i klinisk praksis. Som en konsekvens av et underestimat av kjernetemperatur, vil mange pasienter undertriageres og potensielt kunne gå glipp av viktig behandling når man benytter seg av temporalmålet i akuttmottaket.

Styrker og begrensninger

Metoden for innsamling av data er tilnærmet lik klinisk praksis i akuttmottaket, og det har ikke vært seleksjon av pasienter etter spesifikke kriterier. Dette har gitt et tilnærmet randomisert utvalg fra populasjonen, og gitt et bilde av temperaturmålernes egenskaper i klinisk virkelighet. Alle målingene er utført av to studenter etter produsentens brukerveiledning. Dette har sikret en gjennomgående god kvalitet av målingenes tekniske utførelse, og bidrar til å eliminere systematiske målefeil og produserer et tallmateriale som egner seg for vurdering av måleapparatets tekniske egenskaper. På den andre siden vil det faktum at målingene er utført av to studenter, og ikke de ansatte triagesykepleierne, være en faktor som avviker fra klinisk hverdag. Studien ville også blitt styrket ved å inkludere flere deltagere, og helst en større andel febrile pasienter. Det er få ekstremverdier i utvalget. Dette kan skyldes eksklusjonskriteriene, som forhindrer de sykeste pasientene fra å delta i studien.

Konklusjon

Det gjennomsnittlige avviket mellom temperatur målt temporalt og rektalt er ikke klinisk signifikant. Temporalmålet gir et godt temperaturestimat hos afebrile pasienter, men er noe dårligere hos de febrile, der det har en tendens til å underestimere temperaturen. Dette gir temporalmålet en relativt lav sensibilitet for å oppdage feber, og kan resultere i feildiagnostikk og utilstrekkelig behandling. Det er viktig at klinikere er klar over apparatets begrensninger for å unngå dette. Apparatet har høy reliabilitet, er svært praktisk i bruk og skånsomt for pasienten, men dette må ikke gå på bekostning av pasientsikkerheten.

Interessekonflikter
Ingen.

Referanser

1. Hoffman, R.J., et al., *Comparison of temporal artery thermometry and rectal thermometry in febrile pediatric emergency department patients.* Pediatr Emerg Care, 2013. **29**(3): p. 301-4.
2. Barnason, S., et al., *Emergency nursing resource: non-invasive temperature measurement in the emergency department.* J Emerg Nurs, 2012. **38**(6): p. 523-30.
3. Hooper, V.D. and J.O. Andrews, *Accuracy of noninvasive core temperature measurement in acutely ill adults: the state of the science.* Biol Res Nurs, 2006. **8**(1): p. 24-34.
4. Jensen, B.N., et al., *Accuracy of digital tympanic, oral, axillary, and rectal thermometers compared with standard rectal mercury thermometers.* Eur J Surg, 2000. **166**(11): p. 848-51.
5. Furlong, D., et al., *Comparison of temporal to pulmonary artery temperature in febrile patients.* Dimens Crit Care Nurs, 2015. **34**(1): p. 47-52.
6. Fulbrook, P., *Core temperature measurement in adults: a literature review.* J Adv Nurs, 1993. **18**(9): p. 1451-60.
7. Pompei, M., *Temperature Assessment via the Temporal Artery: Validation of a New Method.* 1999(Validation of a new method): p. 40.
8. Reynolds, M., et al., *Are temporal artery temperatures accurate enough to replace rectal temperature measurement in pediatric ED patients?* J Emerg Nurs, 2014. **40**(1): p. 46-50.
9. Nordas, T.G., S. Leiren, and K.S. Hansen, *[Can ear temperature measurement be used in a hospital?].* Tidsskr Nor Laegeforen, 2005. **125**(20): p. 2763-5.
10. Holte, T.V., P.O; Elvsaas, I; Norderhaug, I, *Diagnostisk nøyaktighet av øre-, munnhule-, armhule- og pannetermometer sammenliknet med rektaltermometer for å identifisere feber hos voksne pasienter innlagt i sykehus eller sykehjem.* 2009: p. 49.
11. Mackowiak, P.A., *Concepts of fever.* Arch Intern Med, 1998. **158**(17): p. 1870-81.
12. Kim, T.K., *T test as a parametric statistic.* Korean J Anesthesiol, 2015. **68**(6): p. 540-6.
13. Bland, J.M. and D.G. Altman, *Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet, 1986. **1**(8476): p. 307-10.
14. Dybwik, K. and E.W. Nielsen, *[Infrared temporal thermometry].* Tidsskr Nor Laegeforen, 2003. **123**(21): p. 3025-6.