

Smertescoringsverktøy for barn

*Hva finnes av evidens, validitet og bruksområder
for smertescoringsverktøy for barn
i alderen 1–12 år?*

Camilla Gustavsen



Litteraturstudie ved det medisinske fakultet

UNIVERSITETET I OSLO

Dato

05.02.2020

Copyright Forfatter

År: 2020

Tittel: Smertescoringsverktøy for barn

Forfatter: Camilla Gustavsen

<http://www.duo.uio.no>

Trykk: Reprosentralen, Universitetet i Oslo

1. Abstract

Aim: To examine what is known about evidence and validity of today's pain assessment tools for children 1-12 years old, and about when to use these tools.

Background: Assessment of pain in the pediatric patient is difficult. Children do not have fully developed cognitive and verbal skills and therefore express pain differently than adults. Still, accurate and good pain assessment in children is important because pain causes severe stress reactions, with potential adverse effects both in the short and long term. Pain may also have a huge impact on the patient's quality of life. Experiences gained early in life may affect help-seeking behavior and use of health care resources throughout life. Children, especially those who are preverbal or nonverbal, are at high risk for having their pain overlooked and its treatment subsequently disregarded. Various pain assessment tools have been developed to help healthcare professionals detect, interpret, and treat children's pain.

Methods: PubMed was searched 24th January 2019 for papers on development and validation of pediatric pain assessment tools. Both systematic reviews and primary studies were included. Different combinations of search terms were used to obtain as many relevant papers as possible. All articles detected were initially evaluated based on their title and abstract. Papers appearing relevant were subsequently evaluated by reading their full text. Some papers found outside the PubMed search were included, as were information from relevant and reliable websites.

Results: Of 1155 articles identified in the primary search, 63 were included in this review. Twenty more were included from other sources and relevant websites. Eight articles and websites were included exclusively because they contained relevant information about how to evaluate pain assessment tools. Pain evaluation tools included in this study were self-assessment tools (Oucher scale, WBFPRS, FPS, FPS-R, CAS, RPS, NRS, VAS) and observational tools (FLACC, CHEOPS, CHIPPS, TPPPS, OPS, PEEPS, M-PEPPS, BOPS, AHTPS, EVENDOL).

Conclusions: Self-reporting tools are often advocated as the primary method for measuring pain intensity in children, but there are disagreements about at what ages children are able to use them. Observational pain tools should be used if the child is unable to self-report. There are many pain assessment tools available, but more research is needed to better define when, for which groups, and for what type of pain the various pain assessment tools should be used.

2. Innholdsfortegnelse

1. Abstract	2
2. Innholdsfortegnelse	4
3. Forkortelser brukt i oppgaven	5
4. Innledning	6
4.1 Problemstilling og valg av oppgave	6
4.2 Utvalg og avgrensning av oppgaven.....	7
4.3 Aldersutvikling i smerteforståelse.....	8
4.4 Definisjon av smerte.....	9
4.5 Inndeling av smerte	9
4.6 Valg av verktøy for smerteevaluering.....	10
4.7 Selvrapporteringsverktøy for smerte.....	12
4.8 Observasjonelle eller objektive skalaer.....	14
4.9 Smertelokaliseringsverktøy.....	14
4.10 Evaluering av smertescoringsverktøy.....	14
5. Metode	20
6. Resultater	22
6.1 Flytdiagram for inkluderte og ekskluderte studier.....	22
6.2 Selvrapporteringsverktøy	23
6.2.1 Smerteverktøy basert på ansikter	23
6.2.2 Smerteverktøy basert på numeriske skalaer og farger	30
6.2.3 Ansiktsskalaer versus numeriske skalaer	36
6.3.1 Sammenligning av de ulike observasjonsskalaene	50
7. Diskusjon	53
7.1 Selvrapporteringsverktøy	54
7.1.1 Smerteverktøy basert på ansikter	54
7.1.2 Smerteverktøy basert numeriske verdier og farger.....	59
7.2 Observasjonelle eller objektive smerteverktøy.....	61
7.3 Generelt om smerteverktøy.....	64
8. Oppsummering og konklusjon	67
9. Litteratur	69
10. Appendix	75
Søkestreng i PubMed.....	75

3. Forkortelser brukt i oppgaven

AHTPS	Alder Hey pediatric triage score
BOPS	Behavioral Observational pain scale
CAS	Color analog scale
CHEOPS	Children's Hospital of Eastern Ontario Pain scale
CHIPPS	Children's and infant's post-operative pain scale
CMFS	Child Medical Fear scale
DPC	Derbyshire Children's Hospital paediatric pain chart
EVENDOL	Evaluation Enfant Douleur
FLACC	Face, Legs, Activity, Cry, Consolability scale
FPS	Faces pain scale
FPS-R	Faces pain scale-revised
M-PEPPS	Modified-Preverbal, Earlyverbal Pediatric pain scale
NRS	Numeric rating scale
OPS	Objective pain scale
PEPPS	Preverbal, Early-verbal Pediatric pain scale
TPPPS	The Toddler-Preschooler postoperative pain scale
VAS	Visual analog scale
WBFPRS	Wong-Baker Faces pain rating scale

4. Innledning

4.1 Problemstilling og valg av oppgave

Jeg har for denne prosjektoppgaven valgt å gjøre en litteraturstudie hvor jeg skal ta for meg smertescoringsverktøy for barn og disse verktøyenes evidens, validitet og bruksområder.

Smertepåvirkede pasienter er noe en lege møter i de fleste av sine konsultasjoner, og smerter er en vanlig årsak til at folk oppsøker lege. Smerte kan være svært invalidiserende og ha stor innvirkning på pasientens livskvalitet. Det å få et inntrykk av pasientens smertebilde og ta pasientens smerter på alvor er viktig for å skape god pasient-legerelasjon og skape tillitt. Smertefølelsen kan være vanskelig å sette ord på både for voksne og barn, men kanskje spesielt for barn, siden mange ikke har utviklet en full forståelse av smertesymptomet eller klarer å uttrykke smerten verbalt. Smertesymptomet har vært forsket på i lang tid, og det har vært vanskelig for leger å definere dette helt klart for alle aldersgrupper. Man skal ikke lenger tilbake enn på 1970 tallet, hvor leger trodde at barn under 1 år ikke opplevde smerte, og at denne sansen ikke var fullstendig utviklet før barnet var 8 år gammelt. I dag vet vi at dette er feil. Det stilles til og med spørsmål ved om langvarig smerte hos for tidlig fødte barn kan føre til endret hjerneutvikling og varige atferdsforstyrrelser, og kanskje til og med gi permanente skader og celledød i hvit substans i hjernen (1, 2).

Smerte skaper en alvorlig stressituasjon i kroppen, og fører til massiv frigjøring av stresshormoner som blant annet adrenalin, noradrenalin og kortisol. Disse kan ha ugunstige effekter på kroppen, både på kort og lang sikt. Forskning har vist at ubehandlet smerte i barndommen kan ha mange langvarige og negative senvirkninger, inkludert økt følsomhet for smertestimuli, endringer i immunapparatet, endret nevrofysiologi, endrede holdninger til smerte, og mindre oppsøking av helsevesenet ved smertetilstander senere i livet. Studier har vist at selv nyfødte barn har minner etter smertefulle hendelser. Disse minnene har vist seg å føre til større smerte under senere medisinske prosedyrer, og de kan føre til utvikling av tilstander som «nålfobi». Slike fobier kan føre til ekstreme fysiologiske reaksjoner på medisinske prosedyrer og kronisk unngåelse av medisinsk behandling senere i livet (3). Det er derfor høyest relevant at de smerteverktøyene som brukes er godt nok validert og grundig evaluert, slik at de på best mulig måte kan gi legen et riktig og pålitelig inntrykk av pasientens smertebilde.

Hvordan og hvorvidt man måler smerter hos barn varierer av to hovedårsaker. For det første må verktøyet for vurdering av smerter hos barn ta hensyn til store utviklingsspesifikke forskjeller, på grunn av barns potensielt begrensede evner i verbal kommunikasjon og assosiativ tenkning. For det andre viser det seg at selv om organisasjonen *International Association of the Study of Pain* (IASP) har kommet med anbefalinger om å regelmessig kartlegge smerte blir dette i praksis oftest ikke gjennomført i mange kliniske settinger, til tross for gode hensikter og mange publiserte skalaer. Smertescoringsverktøy blir altså vanligvis ikke tatt i bruk i smertekartlegging av pediatriske pasienter (4). Et kritisk argument knyttet til at skalaene ikke blir tatt i bruk er at det finnes lite dokumentasjon rundt deres validitet og pålitelighet, spesielt knyttet til bruk for ulike alderstrinn (5).

Konsekvensen av å ikke gjøre regelmessige smertevurderinger er at barna ofte underbehandles. For å overvinne noen av utfordringene ved å vurdere barns smerte kreves utvikling av pålitelige, gyldige og aldersmessig tilpassede smertescoringsverktøy. Slike verktøy må finnes både for akutt smerte, hvor evaluering bør bli gjort flere ganger om dagen, og for kronisk smerte, hvor vurderingen kan gjøres sjeldnere, for eksempel en gang om dagen. For at verktøyene skal kunne tas i bruk og være effektive er det viktig at de er enkle, praktiske og nyttige, og at bruken ikke medfører for stor belastning på barnet, omsorgspersoner, eller helsepersonell. Forenkling av smertemåling kan ha flere fordeler, blant annet a) bedret smertelindring hos barn, b) redusert arbeidsbelastning for sykepleiere og annet helsepersonell som følger opp pasienten og c) opprettelse av et felles «smertespråk» for å lette kommunikasjonen om smerte i ulike settinger (4). Med denne oppgaven ønsker jeg å sammenstille informasjonen vi i dag har om smertescoringsverktøy for barn, og vurdere hva slags evidens som finnes for verktøyene som brukes.

4.2 Utvalg og avgrensning av oppgaven

Jeg har i denne oppgaven valgt å se på generelle kartleggingsverktøy for smerteintensitet, dvs. verktøy som ikke er utviklet for spesifikke diagnoser eller tilstander. Smertescoringsverktøyene er kategorisert enten som *selvrapporteringsverktøy* eller som *observasjonsverktøy* (såkalte «objektive» skalaer). Observasjonsverktøy finnes utviklet for helsepersonell og for omsorgspersoner. Rene smertelokaliseringsverktøy er ikke spesifikt diskutert i oppgaven, selv om disse ofte brukes som supplement i smertevurdering.

Jeg har valgt å begrense utvalget av smerteverktøy som skal vurderes til de som er utviklet for barn i 1-12 års alder. Denne aldersgruppen er valgt fordi det i denne perioden skjer størst kognitiv utvikling og det er størst endring mellom alderstrinnene.

4.3 Aldersutvikling i smerteforståelse

Når barn er 1 år begynner de å utvikle enstavelsesord. De har også et tydeligere og mer forståelig kroppsspråk enn barn under 1 år. Rundt 2 års alder kan de fleste barn kommunisere at de har smerter ved å bruke ord og nonverbale tegn som de har lært av foreldrene. Ved 3-4 års alder har de fleste barn gradvis lært å differensiere smerte og kan ofte beskrive fire nivåer av smerteintensitet (ingen smerte, litt smerte, mer smerte, mye smerte). Ved 5 års alder kan de fleste barn kommunisere smerten mer avansert og er i stand til å bruke selvrapporтерingsverktøy som ansiktsskalaer, selv om de ikke har de samme kommunikative ferdighetene som 6-åringer (6). Førskolealder er en periode med rask kognitiv utvikling, som er nødvendig for effektiv selvrapportering av smerte. For å bruke selvrapporтерingsverktøy trenger barnet å utvikle et språk som kan brukes til å beskrive smerte, og en evne til gjenkjenning av symboler og ansiktsuttrykk.

Forskning har delt inn barns forståelse av smerte i henhold til aldersgrupper (7):

- 2-7 år: Erfarer smerte stort sett som en fysisk forekomst. Forstår smerte som et angrep på dem som følge av egosentrisitet eller noe de har gjort feil. Har problemer med å knytte årsak og virkning til smerteopplevelsen. Kobler smerte til straff for ondskap eller dårlige tanker. Kobler smertereduksjon til «magiske» effekter av smertestillende midler. Barna kan vise aggresjon mot klinikker som utfører smertefulle inngrep, og forstår ikke at smertefulle prosedyrer kan ha gunstige effekter på helsen.
- 7-11 år: Kan identifisere den spesifikke kroppsdelen eller indre organet som er påvirket av smerte. Kan ha frykt for død eller kroppslig skade som følge av smerteopplevelsen. Det er nødvendig som kliniker å gi barnet tid til å beskrive smerten, og barnet trenger klare forklaringer på deres smerte og forklaringer rundt frykt for død og kroppslige skader.
- 12 år og eldre: Kan rasjonalisere betydningen av smerte på mer komplekse måter. Kan likevel fortsatt forestille seg en ondskapsfull årsak til sin smerte. De kan ha problemer med å opprettholde en konsekvent og moden respons på smerten og bruke nyttige mestringsstrategier. Barnet trenger tydelige forklaringer på sin sykdomstilstand, mulighet til å diskutere sin frykt, og påminnelse rundt gode mestringsstrategier.

Barn over 12 år faller utenfor utvalget i oppgaven. Disse barna har oftest god forståelse og gode evner til å kommunisere, slik at man ofte vil benytte smerteverktøy som også brukes av voksne. Oppgaven diskuterer heller ikke egne smertescoringsverktøy for nyfødte og prematurfødte barn (8). Videre eksklusjonskriterier står oppført i metodedelene.

4.4 Definisjon av smerte

Smerte har vært definert slik: «*Smerte er en ubehagelig sensorisk og emosjonell opplevelse som følge av faktisk eller potensiell vevsødeleggelse. Smerte er alltid subjektiv. Smerte er et sammensatt fenomen, som influeres av fysiske, psykiske, sosiale, kulturelle og åndelige/eksistensielle forhold*» (9, 10). Smertevurdering tar altså utgangspunkt i den biopsykososiale modell (3), som gjør det nødvendig å se hele mennesket når man skal evaluere smerte. Definisjonen i seg selv sier mye om kompleksiteten i smertesymptomet. Nettopp det at smerten alltid er subjektiv gjør det vanskelig å finne et standardisert smerteverktøy som kan brukes av alle. Kompleksiteten blir enda større når man som lege skal tolke svarene disse skalaene gir. Forståelsen rundt smerte og de fysiske, psykiske, sosiale og kulturelle faktorene rundt smerten kan gjøre det komplisert å skille mellom hva som er den «opprinnelige vevs-smerten» og hva som skyldes andre faktorer som kan påvirke smertebildet.

Smertevurdering kan forvanskas av at redsel og frykt både ko-eksisterer med og påvirker smerter. Er det for eksempel slik at barnet faktisk har store smerter etter et operativt inngrep, eller kan det være redsel barnet gir uttrykk for? Denne vurderingen er ekstra utfordrende hos barn fordi barn er kognitivt umodne, og de minste barna har heller ikke evnen til å uttrykke smertefølelsen verbalt. Smerte og angst/frykt er to av de vanskeligste følelsene å skille, og det er også vist at disse to følelsene påvirker hverandre. Jo større angst barnet har, jo mer smerte vil det rapportere (11). Det er derfor viktig at man tar høyde for dette når man bruker smerteevalueringsverktøy.

4.5 Inndeling av smerte

Smerte kan være akutt, tilbakevendende, kronisk, eller en kobinasjon av akutt og kronisk. Om smerten er akutt eller kronisk er noe av det første helsepersonell må finne ut av når de møter et barn med smerter. Smertetypen kan ha avgjørende betydning for hvilket smerteverktøy man velger å benytte (7).

Akutt smerte er typisk en advarsel på sykdom, vevsødeleggelse, eller en truende tilstand for kroppen. Det er en nociceptiv form for smerte, og er koblet til en organisk årsak. Akutt smerte har definisjonsmessig relativt kort varighet, og ved varighet lengre enn 3 mnd bør man ikke lenger bruke begrepet «akutt». Akutt smerte kan være assosiert med medisinske prosedyrer (f.eks. kirurgi, injeksjoner), skade (slag, blåmerker, frakturer), akutt sykdom (meningitt, tonsillitt), og forverring av kroniske tilstander (artritt, ulcerøs colitt) (7, 10).

Kronisk smerte er et begrep som brukes for å beskrive en vedvarende eller tilbakevendende smertestilstand. Dette er typisk definert som en smerte som varer lengre enn forventet, mer enn 3 mnd, eller er en tilbakevendende smerte som oppstår minst 3 ganger over en periode på 3 mnd, som for eksempel hodepine (10). Europeiske studier har funnet at rundt 25% av barn har en eller annen form for kroniske smerter, der de vanligste smertetilstandene er hodepine, abdominale smerter og smerter i ekstremiteter. Det er vist en klar sammenheng med alder, der vedvarende smerte er vanligst hos jenter i alderen 12-14 år (7).

Smerten kan også deles inn etter **hvilke nerveceller smertesignalet sendes gjennom**

- **Nociceptiv** smerte er forårsaket av stimulering av nociceptorer som et resultat på skade eller inflammasjon. Den er delt inn i **somatisk** smerte som oppfattes via reseptorer i hud, blodårer, skjelettmuskel og benvev, og **visceral** smerte som oppfattes via reseptorer i indre organer. Somatisk smerte er tydelig lokalisert og kjennes som en skarp, vond, klemmende, stikkende eller knusende smerte. Visceral smerte er mer diffust lokalisert og beskrives ofte som stump (8).
- **Nevropatisk** smerte er forårsaket av stimulering av eller unormal funksjon i ødelagte sensoriske nerveceller. Dette kan forårsakes av kompresjon, transeksjon, infiltrasjon, iskemi eller metabolsk skade av nervevev. Smerten beskrives gjerne som brennende, skytende, elektisk eller ilende (5).

4.6 Valg av verktøy for smerteevaluering

Smerte hos barn blir i mange tilfeller underestimert og underbehandlet grunnet for dårlige kartleggingsverktøy (3). Rapporteringen og derfor evalueringen av smerte påvirkes av barnets kognitive funksjon, klinisk kontekst, og smertens typologi. Med økende alder vil barn etterhvert klare å identifisere et smertestimuli, lokalisere det, kvantifisere det, og tilslutt uttrykke smerten verbalt. Smerte induserer forandringer i oppførsel og fysiologiske- og metabolske forandringer.

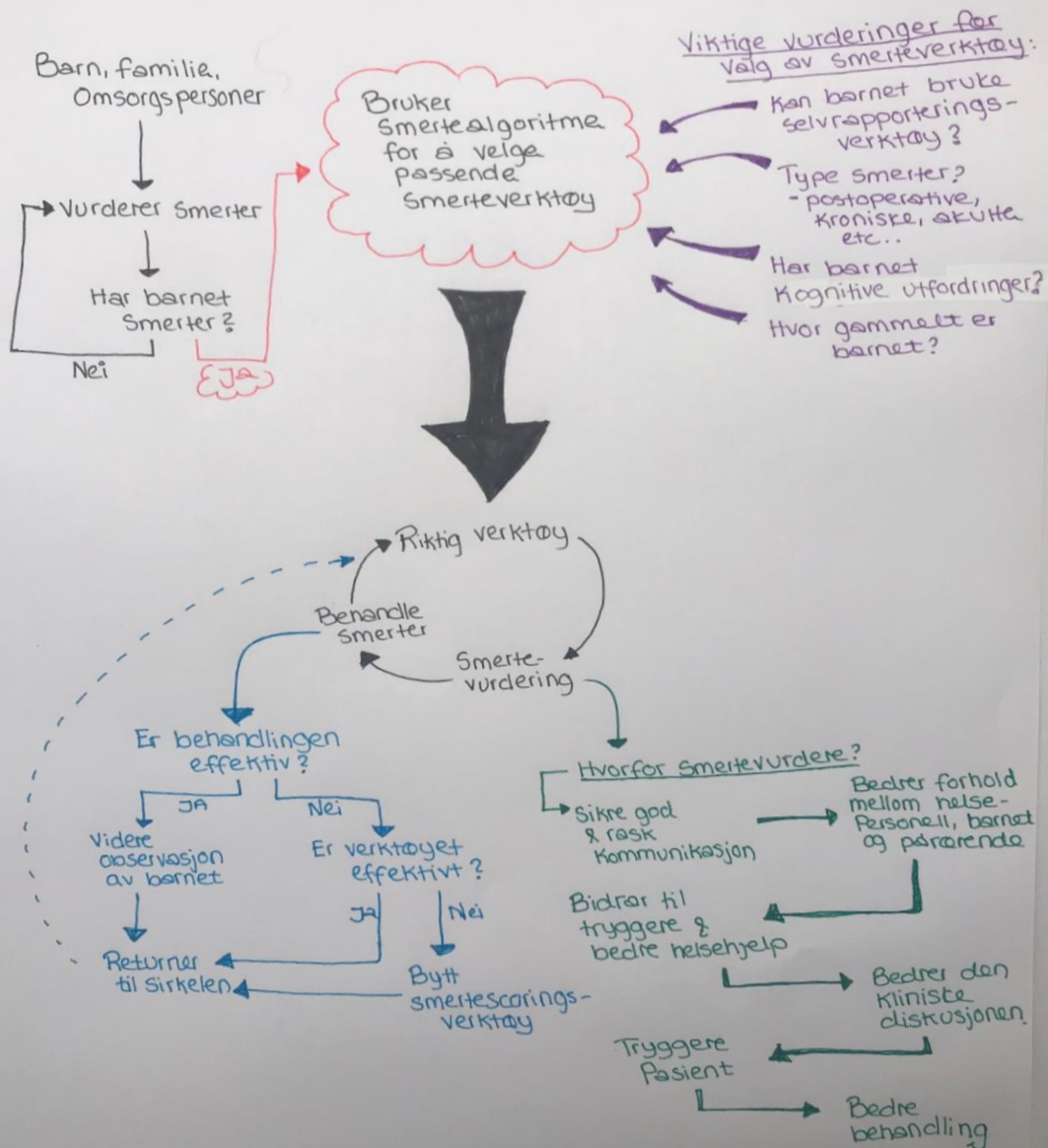
Det er noe uenighet om i hvilke aldersgrupper man kan benytte selvrapporteringsverktøy for smertegrad. Noen mener at det er gullstandard å benytte selvrapporteringsverktøy for barn >6 år, da disse barna som oftest klarer å uttrykke smerte verbalt. Observasjonelle, eller «objektive», skalaer bør brukes til barn <6 år og andre barn som pga. sitt utviklingsnivå uttrykker smerter utelukkende nonverbalt, dvs. gjennom kroppsspråk som ansiktsuttrykk eller kroppsbvegelser. Andre studier konkluderer imidlertid med at man hos yngre barn (3-8 år) bør benytte visuelle–analoge skalaer som f.eks ansikts-skalaen, og at man hos barn >8 år kan benytte numerisk–analoge skalaer (2, 8).

Smertevurdering både via selvrapportert smerte og gjennom observert smerteadferd kan kombineres med fysiologiske målinger av f.eks blodtrykk, puls og respirasjonsfrekvens, SpO₂, og svetting. Disse fysiologiske målingene har imidlertid noe begrenset verdi som smertemål, siden man her benytter det faktum at smerte er en stressor som aktiverer kompensatoriske mekanismer i det autonome nervesystem. Aktivering av det sympatiske nervesystem gir fysiologiske endringer som pulsøkning, svetting, og dilatasjon av pupiller, representative for «fight-or-flight» responsen. Slike fysiologiske målinger er sensitive for endringer i smerteintensitet, men de kan også representere frykt, angst, eller det fysiske «arbeidet» ved gråt. Fysiologiske målinger mangler derfor spesifisitet i forhold til måling av smerte (2, 3, 12).

Mange ulike smerteverktøy ligner på hverandre og er bygget opp på samme måte. De observasjonelle («objektive») skalaene tar utgangspunkt i atferdsendring og fysiologisk respons, mens de subjektive skalaene er selvrapporterende numeriske skalaer som for barn er visualisert i form av for eksempel ansikter eller fargeskalaer som skal gjøre graderingen mer forståelig. Det har blitt utviklet algortimer for hvordan og når helsepersonell skal gjennomføre smertescoring hos barn, for å gjøre det enklere for helsepersonell å vite når de skal ta i bruk verktøyet, og hva de skal vurdere videre. Bildet under er utarbeidet fra flytdiagrammet fra The Royal College of Nursing, UK (Figur 1) (7).

Nedenfor diskuteres noen av de vanligste smertescoringsverktøyene gruppevis, basert på litteratursøk. Ytterligere beskrivelse, evaluering og bruksområder presenteres i resultat- og diskusjonsdelen av oppgaven.

SMERTEALGORITME



4.7 Selvrapporтерingsverktøy for smerte

Det anses ønskelig å benytte selvrapporтерingsverktøy når dette er mulig utfra pasientens utviklingsnivå og sykdomstilstand. Selvrapportering anses mest pålitelig fordi det er bred enighet om at smerte først og fremst må forstås som en subjektiv opplevelse, som angitt i definisjonen av smerte vedtatt av internasjonale studier av smerte (13, 14). Det finnes i dag

over 30 ulike selvrapporteringsverktøy for smerte, men for bare seks av disse finnes veldokumenterte evalueringer som angir pålitelighet, gyldighet, og klinisk nytte hos barn og ungdom med akutt og kronisk smerte (10).

Selv om selvrapporteringsverktøy er å foretrekke og regnes som gullstandard for å vurdere smerte kan det være begrensninger ved bruk av disse skalaene på grunn av barns kognitive, sosiale og kommunikasjonsferdigheter (15). Rapporteringen kan også påvirkes av konteksten, som for eksempel hvem som stiller spørsmålene om barnets smerte, og av barnets forventninger rundt konsekvensene av rapporteringen. For eksempel kan barnet underrapportere smerte i frykt for hva som kan skje hvis de forteller at de har veldig vondt (f.eks. sprøytestikk), eller motsatt hvis de tror det er fordelaktig å rapportere høy score (f.eks. at de får være hjemme fra skolen). Det er derfor viktig når man bruker disse verktøyene at man gjør det med kunnskap om den sosiale konteksten og barnets motiver og forventninger (16, 17). I tillegg endres barns forståelse av selvrapportering seg dramatisk med økende alder, spesielt fra 3 til 7 års alder (16).

Det er i dag ingen dokumentasjon for at én enkelt smerteintensitetskala er pålitelig nok til å måle alle typer smerter hos pediatriske pasienter (10, 11, 18). En systematisk oversikt gjort av blant annet Deborah Tomlinson og Carl L von Baeyer (19) fant at de vanligste brukte selvrapporteringsverktøyene for barn er basert på bilder av ansiktsuttrykk. I denne gruppen er det fire ulike verktøy som har gjennomgått nødvendig evaluering for både akutt og sykdomsrelatert smerte: 1) Faces pain scale (FPS), 2) Faces pain scale-revised (FPS-R), 3) Oucher pain scale, og 4) Wong-Baker Faces pain rating scale (WBFPRS). Disse skalaene er spesielt å foretrekke for barn i alderen 3-12 år.

Det er også utviklet ulike selvrapporteringskalaer som baserer seg på numeriske verdier. De vanligste av disse er 1) Color analog scale (CAS), 2) Numeric rating scale (NRS), og 3) Visual analog scale (VAS). CAS er en vertikal skala som bruker fargeintensitet for å angi smerte; fargen rødt er knyttet til økende smerte (20). VAS og NRS er visuelt fremstilt som en horisontal linje med tall fra 0-5, eller 0-10. Økende tall representerer økende smerte. Forskjellen mellom VAS og NRS er at NRS også kan brukes uten visuell fremstilling (21). Disse skalaene egner seg for barn over 8 år som har en forståelse av numeriske verdier. Det er viktig å sikre seg at barnet har evnen til å rapportere smerte på denne måten før deres analysesvar blir ansett som gyldige. Dette gjelder spesielt i aldersgruppen 3-7 år. Videre må barnets aktuelle situasjon tas i betraktning. Når et barn er i ferd med å våkne opp fra anestesi etter et operativt inngrep er det diskutabelt om svar fra selvrapporteringsverktøy kan ansees

som gyldige. Faktorer som rastløshet, spenning, agitasjon eller sedasjon påvirke selvrappor-
teringen, og i denne perioden kan derfor være nødvendig å bruke et observasjonsverktøy (22).

Det nyeste innenfor smerteverktøy er at det utvikles ulike computerbaserte skalaer som tar
utgangspunkt i de opprinnelige skalaene, men digitaliserer dem slik at man kan bruke dem på
PC eller som en app til mobilen. Målet er å gjøre verktøyene mer brukervennlige, og mer til-
passet hver enkelt bruker (23-26). Man prøver også å rette opp mulige feilkilder ved
papirskalaene (23); dette vil utdypes nærmere i diskusjonsdelen.

4.8 Observasjonelle eller objektive skalaer

Smerteevaluering av nonverbale og preverbale barn kan ikke basere seg på egenrapportering.
Det er derfor utviklet egne scoringsverktøy for disse barna. Disse smerteskalaene er basert på
scoring basert på barnets ansiktsuttrykk, evne til å bli trøstet, nivå av interaksjon, motorisk
respons i ekstremiteter og trunkus, og verbal respons (8). Det finnes svært mange skalaer av
denne typen, men ikke alle er godt dokumentert. De fleste skalaene benytter versjoner av de
samme parameterne, men skalaene kan ha noe ulike cut-off verdier og tillegg eller fravær av
enkeltparametere. Det hyppigst refererte observasjonsverktøyet for smertevurdering av barn
er Face–Legs–Activity–Cry–Consolability (FLACC) -skalaen, mens andre lignende skalaer
er CHEOPS, TPPPS, PEPPS, M-PEPPS OPS, CHIPPS, BOPS, AHTPS (3, 18, 22, 27-29). Se
avsnitt 3.0 Forkortelser

4.9 Smertelokaliseringsverktøy

Smertelokaliseringsverktøy er utviklet til å kartlegge hvor barnet har vondt. Verktøyet består
vanligvis av en skisse av kroppen, og pasienten blir bedt om å fargelegge området (-er) med
smerter. En kjent skala av denne typen er Adolescent and Pediatric Pain Tool (APTT), der
barnet får en skisse av fremre og bakre del av kroppen og fargelegger området (-er) med
smerter. Lokaliseringsverktøy kan brukes som et supplement til selvrapporтерingsverktøy, for
å få en mer helhetlig vurdering av smerten (30).

4.10 Evaluering av smertescoringsverktøy

Om en smerteskala er av klinisk verdi eller ikke, avhenger først og fremst av skalaens
validitet og reliabilitet (31). For å undersøke dette kreves det forståelse av de forskjellige
komponentene innenfor disse begrepene, og metoder for hvordan disse kan testes.

VALIDITET – GYLDIGHET	Sier noe om undersøkelsen måler det som er tenkt til og måles. For eksempel: undersøker smerteverktøyet smerte, eller måler det andre årsaksforhold som for eksempel frykt. Et av hovedpoengene med eksperimentelle undersøkelser er å trekke slutninger som er valide (32). Man har ulike former for validitet:
a) Innholdsvaliditet Content validity	Er innholdet i testen (de enkelte variabler som f.eks. ansiktsuttrykk, gråt, kroppsspråk osv.) representable for parameteren testen skal måle (smerte)? Ofte forsøker man å besvare spørsmålet ved å fremlegge instrumentet for eksperter innenfor faget. Innholdsvaliditeten er viktig for utviklingen av instrumentet, men gir ikke akseptable bevis for validitet utover dette. Innholdsvaliditet kan tildels vurderes uten formelle statistiske analyser (33).
b) Kriterievaliditet Criterion validity	Er en empirisk validitet som innebærer hvor godt en testvariabel korrelerer med et bestemt kriterium. Kriterievaliditet er et av de mest pålitelige validitetsparametrene. Samtidig validitet (Concurrent validity) er en undergruppe av kriterievaliditet som innebærer at man sammenligner resultatet mellom to instrumenter som måler det samme, på samme tidspunkt. Et eksempel er hvis man undersøker en ny smerteskala og samtidig bruker en skala som er godt utprøvd og tilslutt sammenligner resultatene. Høy overenstemmelse mellom de to skalaene tyder på høy validitet (32).
c) Konstruksjonsvaliditet Construct validity	Uttrykker om et spørreskjema eller en test måler det som er ment å måle. Undersøker om verktøyet oppdager teoretisk forutsigbare endringer i konstruksjonen. En relevant undersøkelse av konstruktvaliditeten er sammenhengen med assosierte fenomener, for eksempel om smerteintensitet synker ved

	administrering av smertestillende (33). Konstruksjonsvaliditeten er selve kjernen i valideringsprosessen fordi andre typer validitet-relaterte bevis støttes av konstruksjonsvaliditeten til testen(34).
d) Indre /intern validitet	Uttrykker om resultatene er korrekte og gyldige for det store utvalget. Randomiserte studier har høyere indre /intern validitet på grunn av randomiseringen og definerede inklusjons- og eksklusjonskriterier. Det er derimot usikkert om resultatene er gyldige for andre pasientgrupper (35).
e) Ytre / ekstern validitet	Angir i hvilken grad resultatene er gyldige under andre betingelser og for andre utvalg enn hva de i utgangspunktet var utviklet for – generaliserbarhet (35).
f) Diskrimineringsvaliditet	Er en undergruppe av konstruksjonsvaliditet. Denne er høy dersom to variabler som antas å ikke være relatert, faktisk ikke er det. Vises ofte ved at det ikke er noen korrelasjon mellom variablene (36).
g) Konvergent validitet	Er en undergruppe av konstruksjonsvaliditet og sier at en test som er designet for å måle en bestemt variabel (f.eks. smerte) faktisk måler det. Oppnås ofte ved hjelp av å undersøke om det er en korrelasjon mellom to målevariabler (36).
h) Known groups validitet	Er en undergruppe av konstruksjonsvaliditet hvor en gruppe barn med kjent grunn til å ha smerter sammenlignes med en gruppe barn som er forventet å ikke ha smerter (37).
i) Umiddelbar validitet Face validity	Sier noe om verktøyet/smerteskalaen virker klinisk fornuftig og gyldig ut fra brukerens subjektive mening, kun basert på inspeksjon av spørsmålene uten at det utføres formelle statistiske analyser (38).
KORRELASJON	Ser på sammenheng mellom variabler. Måles ved hjelp av en korrelasjonskoeffisient, r, på en skala fra -1 til 1.

	<p>$r = -1$ er perfekt negativ korrelasjon, mens $r = 1$ er perfekt positiv korrelasjon. $r=0$ betyr ingen korrelasjon. Korrelasjon kan også plottes ved hjelp av lineær regresjon (36).</p>
<p>RELIABILITET</p>	<p>Sier om det er pålitelighet i målingene som er gjort, og om metodene kan etterprøves av andre og gi tilsvarende resultatet. Reliabiliteten er knyttet til om undersøkelsen representerer den virkelige situasjonen. Testen har høy reliabilitet dersom de tilfeldige feilene er små. Dette sier noe om testens nøyaktighet (33, 39). Bevis på reliabilitet signaliserer til brukeren av testen at resultatet er konsistent og mest sannsynlig til å stole på (40). Ulike måter å estimere reliabiliteten til et måleinstrument er:</p> <ul style="list-style-type: none"> • Test-retest: Dette er den vanligste måten å undersøke reliabilitet på. Man regner ut korrelasjonen mellom to målinger av samme test på forskjellige tidspunkt. En ulempe med denne testen er at personen kan huske hva de svarte i første måling. En annen ulempe med denne formen for reliabilitets-testing er at den den egner seg dårlig for egenskaper som ikke er stabile (f.eks. følelser eller symptomer), siden dette endrer seg over tid (38, 40, 41). • Alternativ form: lager to versjoner av samme test, som er lik mht. innhold, men som ikke inneholder identiske spørsmål. En ulempe med denne skalaen er at man må forutsette at de to versjonene er like. En lav korrelasjon kan bety enten at skalaen er upålitelig eller at de to skalaene ikke er ekvivalente (38, 40). • Indre konsistens: Tar utgangspunkt i om flere variabler måler samme begrep (i dette tilfellet smerte). Brukes for å måle reliabiliteten av en skala som

	<p>består av flere variabler, som tilslutt summeres og gir en totalscore (38). Indre konsistens måles ved Cronbach's alfa koeffisient (α). Denne verdien er et mål på hvor sterk statistisk sammenheng det er mellom alle spørsmål/variabler som måler samme begrep. Vanlige krav er en indre konsistens på minimum 0,70. Høy inter-item korrelasjon = god reliabilitet. Det er viktig å være oppmerksom på at denne verdien påvirkes av antall testledd. Flere testledd vil gi en høyere Cronbach's alfa koeffisient (32, 41).</p> <ul style="list-style-type: none"> • Split-half: Er et enklere mål på indre konsistens. En test splittes i to like deler. Man administrerer testen én gang og ser på korrelasjonen mellom de to delene. Ulempen med denne testen er at det er ikke sikkert de to halvdelene er like (40, 41).
a) Interrater-reliabilitet	<p>Reliabilitet til bedømmelser gjort av flere observatører. Brukes dersom ulike observatører tolker testresultatet. Bregner korrelasjonen mellom flere observatørers score av de samme testene. Beregnes med Cohen's kappas koeffisient (38).</p>
b) Intrarater-reliabilitet	<p>Reliabiliteten til bedømmelser gjort av kun én observatør. Graden av samsvar mellom ulike testsvar dersom en person gjennomfører den samme testen flere ganger. Hvis testen har en høy intrarater-reliabilitet vil samme observatør få samme testsvar hvis testen gjentas med samme forutsetninger (38).</p>
Statistisk signifikans	<p>Mål på hvor sannsynlig det er at de dataene man har, skyldes tilfeldigheter. Statistisk signifikans er tilstede dersom observasjonene i en vitenskapelig studie er fordelt på en måte som ikke kan antas å skyldes tilfeldige variasjoner i forhold til den oppstilte nullhypotesen.</p>

	<p>Påvisning av statistisk signifikans betyr ikke nødvendigvis at man har en påvist kausalitet (årsaksforhold). Før studien utføres setter man et signifikansnivå. Vanligst er 5%, 1% eller 0,1%. Nivået velges ut fra hvor stor sjanse man ønsker å ta for feilaktig å forkaste en gyldig nullhypotese (39).</p>
<p>Psykometriske egenskaper / Psychometric properties</p>	<p>Konstruksjon og validering av måleinstrumenter og vurdering av om disse instrumentene er reliable og valide. I atferdsmedisin er psykometri oftest måling av individets kunnskap, evne, personlighet og type oppførsel. Måling skjer ofte med spørreskjema. Skjemaet må evalueres før man kan oppgi at de har gode psykometriske egenskaper, dvs. at man mener skalaen er reliabel og valid (37).</p>

Mens reliabilitet viser til instrumentet eller testen, selve målingen, handler validiteten om gyldigheten til resultatet av testen (41). Høy reliabilitet er en forutsetning for at måleinstrumentet skal ha høy validitet: en test med lav reliabilitet kan ikke være valid, da den også vil registrere andre faktorer som ikke har med studien å gjøre. Selv om reliabiliteten er høy, kan testen likevel ha lav validitet; en test som måler angst kan ha høy reliabilitet, men den er ikke valid til å måle smerte. Høy reliabilitet er ikke et bevis på høy validitet, men lav reliabilitet er bevis på lav validitet (32).

En enkelt studie er ikke tilstrekkelig for å fastslå validitet og reliabilitet til en skala. Det trengs flere studier av andre forskere enn kun de opprinnelige forfatterne for å kunne fastslå om skalaen har tilstrekkelig validitet og reliabilitet for og brukes i klinisk sammenheng (31). Validiteten og reliabiliteten til et måleinstrument kan variere mellom ulike grupper. Det er derfor ikke gitt at en smertescore som er utviklet for akutt smerte også kan brukes på kroniske smerter (32).

5. Metode

Denne litteraturstudien hadde som mål å vurdere smertescoringsverktøy for barn, og disse verktøyenes evidens, validitet og bruksområder. Litteratursøket hvor hovedmaterialet i denne oppgaven er hentet fra ble gjort i PubMed den 24.01.19, ved hjelp av bibliotekar ved OUS Rikshospitalet. Søkestrategien ligger vedlagt som Appendix 1.

For å søke frem studier på barn 1–12 år ble det brukt søkekombinasjoner som *child* OR *children* OR *preschool* OR *pediatrics* OR *pediatric*. Søkeordet *infant* ble ikke brukt, da PubMed setter dette søkeordet på studier som inkluderer personer fra fødsel til 23 måneders alder. På samme måte ble PubMed-søkeordene *newborn* (fødsel–1 mnd) og *adolescent* (13–18 år) utelukket. Søkeordet *Child* vil få frem de fleste studier av barn i oppgavens inkluderte aldersgruppe, da PubMed definerer *child* som personer fra fødsel til 18 år. Jeg valgte å søke med MeSH terms da dette gir et bredere søk, fordi man får med alle søkeord som PubMed assosierer med «child».

For å søke frem studier av smertescoringsverktøy kombinerte jeg søkeordene for studier på barn med AND (*pain measurement* OR *pain assessment* OR *pain scale* OR *pain tool*). Her søkte jeg også på MeSH terms for «pain» for å få med de ulike ordene PubMed assosierer med «pain». Jeg søkte også på noen spesifikke smerteskalaer, som for eksempel Comfort scale, NRS, FLACC, PPMP, CEHOPS. Fordi målet var å finne studier som fokuserte på *validering* av verktøyene kombinerte jeg disse søkestrengene med AND (*validation* OR *validate* OR *psychometrics* OR *reliability* OR *reliable*). Jeg begrenset søket til artikler basert på «humans» og språket til artikler på engelsk, norsk og svensk.

Artikler som kom frem av søket ble først vurdert utfra tittel og abstrakt. De artiklene som utfra dette virket mulig relevante ble grundig gått igjennom ved å lese fulltekst. Både systematiske oversikter og originalartikler ble inkludert. De systematiske oversiktene ble mest brukt for å få oversikt og bakgrunnsinformasjon av hva som finnes av ulike smertescoringsverktøy. Fulltekstartiklene ble brukt for å få en dybde i oppgavens resultat og diskusjonsdel. Informasjon fra relevante og pålitelige nettsider er angitt som litteraturhenvisninger.

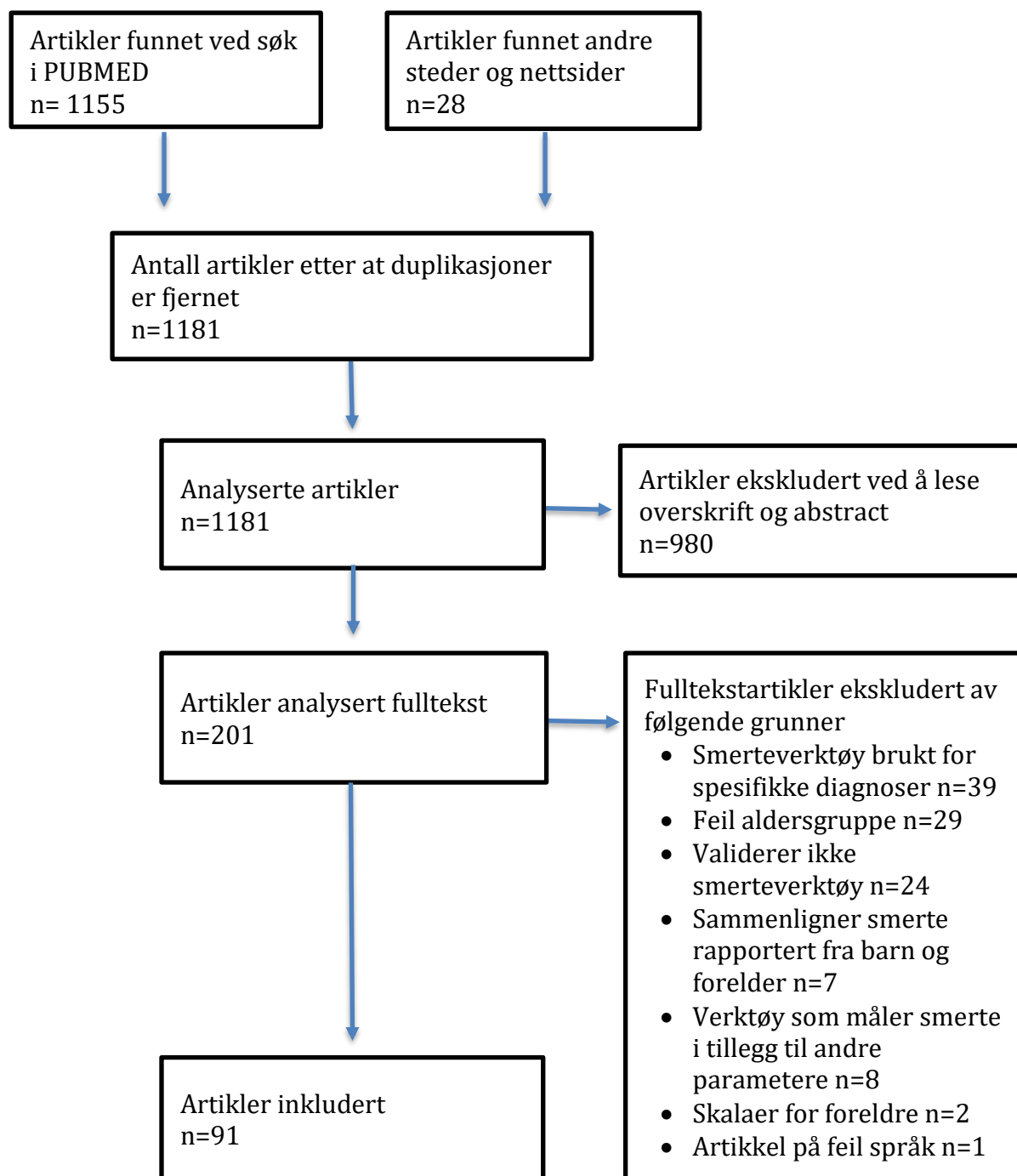
Artikler som ikke tok for seg utvikling eller validering av selve smertescoringsverktøyet ble utelukket. Oppgaven ble avgrenset til studier av verktøy der barn selv rapporterer smerte, eller der kvalifisert helsepersonell vurderer barns smerte ut fra observasjoner («objektive» skalaer). Smerteverktøy hvor foreldre vurderer sine barns smerte ble utelukket. Studier av

verktøy som tar for seg andre parametere i tillegg til smerte, som sedasjon, angst og redsel, ble utelukket.

Jeg valgte å kun inkludere studier av *generelle* smertescoringsverktøy; artikler som omhandlet smerteverktøy for spesifikke diagnoser og tilstander, for eksempel for dental smerte eller onkologiske smerter, ble utelukket. Artikler jeg valgte å ta med i studien tok for seg de generelle smertescoringsverktøyene som brukes av ellers friske barn for eksempel brukt etter postoperative inngrep, prosedyrer osv. Barna hadde ikke annen alvorlig sykdom, som kan komplisere smertebildet. Likevel valgte jeg å inkludere barn med utviklingshemming og mental retardering, ettersom observasjonsverktøy som blant annet FLACC er utviklet nettopp med tanke på denne pasientgruppen. Dette er også en viktig gruppe å inkludere ettersom de i større grad enn andre barn har kroniske smertefulle tilstander, oftere gjennomgår smertefulle prosedyrer, og kan ha nedsatte evner til å uttrykke sitt smertebilde (42).

6. Resultater

6.1 Flytdiagram for inkluderte og ekskluderte studier



Søkestrategien i PubMed ga et treff på 1155 artikler. Etter min vurdering basert på artiklens tittel og abstrakt var 173 artikler mulig relevante. Disse artiklene ble grundig gått igjennom ved å lese fulltekst. Flytdiagram viser årsaker til at artikler ble ekskludert. Sluttresultatet ble 63 artikler som jeg anså som relevante å ta med i studien (2-4, 6, 7, 10-29, 31, 33, 34, 42-76). Totalt antall artikler inkludert i studien er 91, av disse er 20 funnet andre steder som gjennom artikler i søket, eller nettsteder som blant annet Helsebiblioteket (1, 5, 8, 9, 30, 77-91). Åtte artikler og nettsider ble spesifikt funnet for avsnittet «evaluering av smertescoringsverktøy» (32, 35-41).

Hovedutfordringen med litteratursøket var å finne artikler som tok for seg validering og evidens av smertescoringsverktøyene. De fleste artiklene validerte ikke selve verktøyet, men verktøyet ble brukt til å måle smerte etter en prosedyre, eller til å vurdere en behandlingseffekt.

6.2 Selvrapporteringsverktøy

Selvrapporteringsverktøy er som beskrevet tidligere det man helst ønsker å benytte dersom barnet er i stand til å bruke det (13, 14). I dag er det hovedsakelig to ulike grupper av selvrapporteringsverktøy: Smerteverktøy basert på ansikter og smerteverktøy basert på numeriske verdier.

6.2.1 Smerteverktøy basert på ansikter

I søket var det 19 artikler som omhandlet kun ansiktsskala-verktøy (4, 6, 19, 23, 24, 26, 47-51, 58, 61, 63, 70, 71, 74, 79, 81). De fleste artiklene tok for seg FPS/FPSR og WBFPRS (4, 6, 19, 23, 50, 58, 61, 63, 70, 71, 79, 81), mens de resterende handlet om én mindre kjent ansiktsskala (51), computer baserte skalaer (24, 26), og Oucher pain scale (47-49, 74). Flertallet av de 19 artiklene tok for seg mer enn én skala. Carl L. von Baeyer (4, 6, 19, 79) og Judith E. Beyer (47-49) har vært med å utarbeide flere av artiklene om de ulike smertescoringsverktøyene

Ansiktsskalaene består av et kontinuum av tegninger eller bilder av ansikter, hvor hvert bilde representerer ulik smerteintensitet. De har vist seg å være passende for barn da de ikke inneholder vanskelige ord eller abstrakte numeriske verdier (51). De har også vist seg å ha høy korrelasjon med andre selvrapporteringsverktøy, og for å være sensitive for analgetikabehandling og ikke-farmakologiske intervensjoner og at det er en variasjon avhengig av smertestimuli (53). De er populært brukt i kliniske settinger, og kan

gjennomføres raskt og enkelt. Ansiktsskalaer har generelt vært det verktøyet barna selv foretrekker å bruke, når de har fått valget mellom ulike skalaer (19, 51, 59). Det er dessverre konseptuelle og psykometriske uklarheter ved bruk av disse skalaene. For å kunne bruke ansiktsskalaer må barnet ha evne til å gjenkjenne og matche sine indre følelser av smerte med ansiktsuttrykkene gitt i en ansiktsskala. Dessuten kan det være vanskelig å skille akutt smerte fra stress og angst. Det er også problemer med å etablere i hvilken grad ansiktene i en skala sies å danne et ekte kontinuum (61).

Som referert tidligere er det i dag utarbeidet fire ulike skalaer: Oucher pain scale og Wong-Baker Faces Pain Rating Scale (WBFPRS), Faces Pain Scale (FPS), og Faces Pain Scale-Revised (FPS-R), som kan brukes ved akutte og kroniske smerter (19). Nedenfor gis en kort beskrivelse av disse fire ansiktsskalaverktøyene.

Oucher scale: Er en av de eldste og mest testede verktøyene. Skalaen ble utviklet av Beyer i 1980, og er et selvrapporteringsverktøy som hovedsakelig ble laget for hvite barn i alderen 3-12 år (47-49). Verktøyet består av to separate vertikale skalaer; en numerisk skala (0-100) for større barn, og en fotografisk skala for mindre barn. Den fotografiske skalaen består av 6 ulike bilder av et barn som viser ansiktsuttrykk som skal illustrere en skala fra «no hurt» (ansikt 1) til «the biggest hurt you can have» (ansikt 6). Fotografiene ble tatt under faktiske smerteopplevelser på

sykehuset. Skalaen brukes ved at barnet peker ut det ansiktet de mener sin egen smerte representerer, eller et nummer på den numeriske skalaen (dersom barnet er eldre). Tre versjoner er utviklet

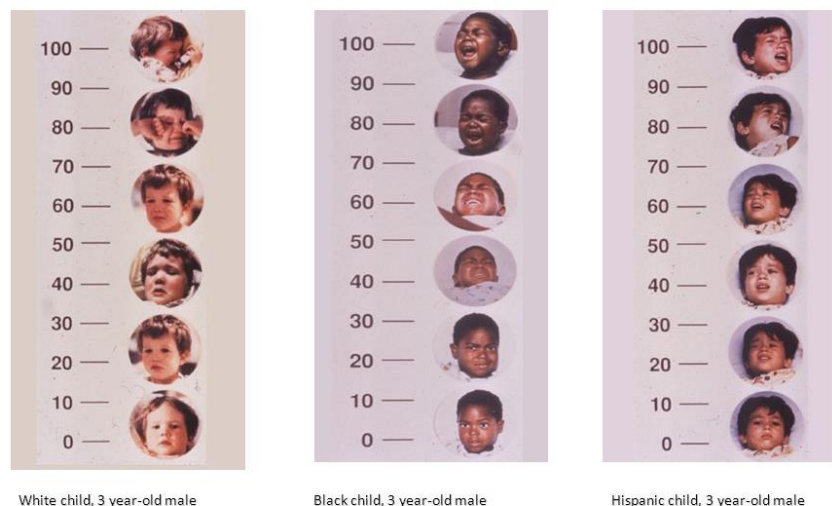
og validert: 1) Afro-amerikansk,

2) Latinamerikansk, og

3) Kaukasisk, hvor det

er bilder av barn i ulike etniske grupper (47-49). Validiteten av den afroamerikanske og

Oucher Photographic / Numeric Pain Scale



Figur 1, 2, 3 (49). Viser de tre validerte versjonene av Oucher pain scale.

latinamerikanske versjonen ble testet av Villarruel Denyes og kolleger (47, 74), fordi de fant at den opprinnelige versjonen ikke var tilstrekkelig valid for barn med annen hudfarge.

I prosessen med å utvikle den opprinnelige Oucher-skalaen gjorde den store størrelsen på de opprinnelige plakatene (14" x 18" plakat) verktøyet noe uhåndterlig og vanskelig å bruke. I 1992 ble derfor størrelsen redusert til 11 inches som gjorde at helsepersonell kunne ta det med seg i lommen. Det ble gjennomført en studie for å vurdere om reduksjon av størrelsen på plakatene endret skalaens pålitelighet (49). Det var totalt 137 barn med i studien, i alderen 3-12 år og fra alle de tre etniske gruppene som skalaen er beregnet for. Barna skulle gjennom en kirurgisk øre-nese-hals-prosedyre. Barna fikk fremstilt både en stor og en liten versjon av skalaen, for å se på samsvar mellom disse to. Den fotografiske skalaen ble brukt av 70 barn i alderen 3-7 år, mens den numeriske skalaen ble brukt av 67 barn i alderen 5-12 år. Av barna som brukte tallskalaen hadde 82,8% den samme verdien på begge skalaene. 80% av barna scoret samme verdi på den store og den lille fotografiske skalaen. De mindre versjonene ble trykt med en av de tidligste modellene av fargelaserskriver. Kopiene var kornete av utseende, og fargene ble forvrengt. Pålitelighetskoeffisientene var tilstrekkelige (0,76 (fotografisk) og 0,94 (numerisk) for «kaukasiske» Oucher). De kornete bildene ble ansett som en mulig grunn til at færre barn hadde samme verdi på den fotografiske modellen, og det ble anslått at med bedre farger kunne påliteligheten forbedres, særlig i den yngre aldersgruppen (49).

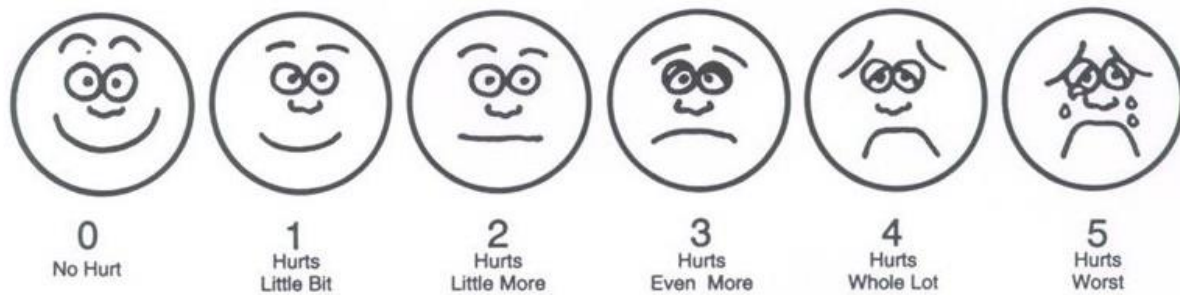
For å avgjøre om et barn var kognitivt modent nok til å bruke Oucher smerteverktøy ble det først gjennomført en geometrisk testing hvor det ble brukt firkanter. Barnet skulle plukke ut hvilken firkant som var den største av en gruppe på seks. Dette ble gjentatt helt til barnet hadde plukket ut alle firkantene. Dersom barnet gjorde denne testen riktig ble de ansett som å være i stand til å bruke Oucher for å vurdere egen smerte. For å avgjøre om barna var i stand til å bruke tallrekken ble de bedt om å telle fra 1-100, og deretter avgjøre hvilket av to tall angitt av testperson som var det største. Barna som kunne utføre disse to abstrakte oppgavene, brukte den numeriske skalaen i Oucher studien. Korrelasjoner varierte fra en middels på 0,875 i unge kaukasiske barn til høy på 0,998 i eldre kaukasiske barn. Disse funnene viser et sterkt, positivt og signifikant forhold mellom scorene på de store og små versjonene av Oucher. Det er også gjort flere studier, av blant annet Jordan-Marsh i 1994, som viser de samme matchende resultatene mellom de store og små versjonene av Oucher skalaen (49).

The Wong-Baker pain Faces scale (WBFPRS): er et verktøy som har blitt foretrukket av barn og foreldre for å rapportere smerte fordi den er enkel å bruke (58). Skalaen ble utviklet

av Wong og Baker i 1988 (63). Den består av 6 svart-hvitt tegneserieansikter som representerer økende grad av smerte bortover på en horisontal linje fra venstre mot høyre. Ansiktet som representerer ingen smerte er smilende, mens ansiktet som representerer mest smerte (nr. 6) er gråtende. Ansiktene skiller seg fra hverandre ved formen på munnen (smilende, naturlig og nedadvendt), ved formen på øyebrynene (nedadkurvende, oppadkurvende og ved å ligge på toppen av øynene) og ved karakteren på øyene (åpne vs. mer lukkede, ingen tårer vs. tårer) (23).

Tegningene ble opprinnelig laget av barn, som ble bedt om å tegne hvordan de selv ville sett ut ved varierende grad av smerte. Ved bruk av verktøyet blir barnet bedt om å peke på ansiktet som angir smerten de føler. Ansiktene er koblet til et nummer 0–1–2–3–4–5, eller eventuelt 0–2–4–6–8–10, som skal representere intensiteten i forhold til en numerisk skala (63). Skalaen sies å være representativ for barn fra 3 år og oppover (5, 63, 81). Det har blitt undersøkt om frykt kan ha noe å si for hvilket ansikt barna velger. I en studie av Gregory Garra (58) og hans forskningsgruppe undersøkte de om WBFPRS hadde god diskrimineringsvaliditet mellom smerte og frykt. Dette fordi det hadde blitt publisert få studier om diskrimineringsvaliditeten til ansiktsskalaer. Forskernes hypotese var at smerteintensiteten rapportert med WBFPRS ville samsvare med en 100mm visual analog smerteskala (VAS), men ikke samsvare med scoringsverdier fra en kjent skala som måler frykt: Child Medical Fear Scale (CMFS). CMFS ble brukt til å vurdere graden av frykt ved medisinske prosedyrer. Skalaen består av spørsmål innen fire kategorier: prosedyre-, omgivelser-, intrapersonell- og interpersonell frykt. Det var totalt 29 utsagn hvor barnet skulle angi 1) ingen frykt, 2) litt frykt 3) stor frykt. Poengene ble slått sammen, for å regne ut den totale poengsummen.

Studien inkluderte 197 barn i alderen 7-12 år. Pasientene ble bedt om å rangere smerten sin både med WBFPRS og VAS. Korrelasjonen mellom WBFPRS og VAS var moderat: 0,63 (95% CI, -0,13 – 0,15). Ansiktet «hurts even more» korrelerte med en 55 mm VAS (på skala 0–100 mm). Korrelasjonen mellom CMFS og WBFPRS/VAS var meget svak: -0,02 (95% CI, -0,16 – 0,12) og 0,01 (95% CI, -0,13 – 0,15). Resultatene var uavhengig av kjønn, alder, smertelokalisasjon og årsak til smerten. Resultatene viste altså at WBFPRS korrelerer moderat med andre smerteskalaer som VAS, men at barn i alderen 7-12 år klarte å skille mellom smerte og frykt (58).



Wong-Baker faces pain rating scale, WBFPRS (23)

Faces pain scale (FPS) og Faces pain scale-revised: FPS og FPS-R er visuelle skalaer med ansikter. **FPS** Skalaen ble utviklet av Bieri et al. (1990) (4, 50). Hensikten med skalaen var å unngå problemene som andre skalaer har, ved at de inkluderer smil og tårer i ansiktene. De ønsket også å utvikle en skala som var lett å bruke, og ikke krevde høye kognitive evner. For å utvikle FPS skalaen ble det inkludert 553 barn, hvorav 195 med median alder 6,8 år og 358 med medianalder 8,7 år (50). FPS besto av 7 ansikter (figur 1 (4)), da forskerne mente at dette ga høyere sensitivitet enn en skala med 5 ansikter, og at en skala med 7 ansikter ville dekke hele dimensjonen av smertegrad. (50) Skalaen ble utviklet gjennom ulike faser, hvor den første fasen var at barna skulle tegne 5 ulike ansikter med ulik grad av smerte. Uttrykk som gikk igjen i flere av ansiktene ble brukt til å utvinne 5 sett med 7 tegninger i hvert sett. Det ble gjennomført en pilotstudie (en mindre versjon av en foreslått studie, utført for å avgrense metodikken til den senere større studien) hvor de innsnevret de 5 settene til 1 sett som delte den høyeste graden av enighet. Etterpå ble barna bedt om å rangere rekkefølgen av ansiktene. Fase 3 tok for seg at det skulle være likt mellom intervallene, og barna plasserte ansikter på et selvvalgt punkt langs en skala. I fase 4 testet de den spontane bruken av skalaen, og barna ble presentert for ansiktene i en tilfeldig rekkefølge og de skulle angi den opprinnelige ansiktsrekkefølgen. Fase 5 ble gjenbrukspåliteligheten testet hvor barna rangerte smertefulle hendelser på dag 1 og etter 2 uker (50).

Resultatene viste at barna i alderen 6 og 8 år hovedsakelig brukte øynene og munnen som de viktigste signalene for endring i smerte. Furing av pannen ble også brukt. Barna brukte få av andre endringsmønstre som tårer, forandringer knyttet til kinn, hake og nese. Det hyppigste uttrykksmønsteret med økende smerte var uregelmessighet i leppene, hvor endende begynte å peke nedover, og progressiv lukking av og rynking av øyebrynene. Tegningene av de 8 år gamle barna hadde en mer konsistent endring i ansiktsmønstre, enn tegningene til de 6 år gamle barna ($p < 0,02$). Ellers var det ingen signifikante forskjeller som følge av alder (50). Da

de undersøkte intervall-lengden mellom ansiktene i fase 3, var dette ikke som de hadde forventet for ansikt 2 og 4, og i mindre grad ansikt 3. Alle barna tok disse ansiktene nærmere hverandre på skalaen enn det forskerne hadde forventet, og de var ingen signifikant forskjell mellom de to aldersgruppene. Likevel hevder forskerne av denne skalaen at den er gyldig og pålitelig, og at den kan betraktes som en intervallskala som passer for barn 6-9 år.

Bieri og hans forskningsgruppe mente likevel at aldersnivået påvirket responsen ved at yngre barn var mer varierende i nøyaktigheten (50). I år 2000 ville Mick Hunter og hans forskningsgruppe (61) bygge videre på dette, og derfor teste om FSP-verktøyet egner seg for barn helt ned i 3 års alder. De inkluderte 135 friske jenter og gutter i alderen 3 ½ -6 ½ år. De delte barna inn i mindre grupper igjen basert på alder og effektberegninger basert på dataene i Bieri (50).

Undersøkelsen ble gjennomført individuelt for hvert enkelt barn, og de undersøkte først barnets forståelse av smertebegrepet, slik at barnets eget «smertespråk» ble brukt. I første fase av studien la de bilder alle de 7 ansiktene ut på et bord, og ba barnet om å rangere ansiktene fra de med minst til mest smerte. I andre fase fikk barnet utdelt 2 ansikter hvor de skulle si hvilket ansikt som uttrykte mest smerte, og i fase 3 fikk de utdelt byggeklosser som skulle representere smerte. 6 klosser representerte ansiktet med mest smerte, og ingen klosser representerte ansiktet med minst/ingen smerte. På de andre ansiktene skulle barnet plassere antall klosser det mente hvert ansikt representerte. I den siste delen av forsøket som tok for seg gjenbrukspåliteligheten fikk barnet fremstilt en rekke smertefulle hendelser (eks skrubb-sår, venepunksjon, å brenne seg, klemming av finger i en dør, lugging osv.), som de skulle rangere bortover på ansiktskalaen. Alle testene ble retestet en uke senere.

Resultatene fra studien viste at alle barna uavhengig av alder klarte å plassere ytterkantene av ansiktens kontinuum med nesten fullstendig nøyaktighet, forutsatt at man forklarer skalaen enkelt og med barnets eget «smertespråk» i fase 1. Det var mer vanskeligheter knyttet til plassering av ansikt 3,4, 5 og 6. Hvor guttene oftest plasserte ansikt 6 feil i rekkefølgen, mens jentene plasserte ansikt 5 feil i rekkefølgen. Utenom dette så man ingen ulikheter mellom kjønnene i de ulike fasene. Det viste seg at barna i aldersgruppen mellom 3 ½ - 4 ½ år hadde større problemer med å velge ut riktig ansikt, når de ble bedt om å velge mellom to stykker, i fase to. I fase tre av testen ble antall klosser som skulle legges på de ulike ansiktene enten over – eller underestimert fra en forventet lineær skala (hvis man tenker seg at det skulle ligge én kloss på ansikt 2, to klosser på ansikt 3, tre klosser på ansikt 4 osv.). Det ble da stilt

spørsmålstegn ved om skalaen kunne brukes som en lineær skala, med konsistente intervaller (61).

Funnene indikerte at et problem med FPS er at det er vanskelig å direkte oversette poengsummene til numeriske tallverdier, slik at helsepersonell kan forstå de bedre. Det ble derfor foreslått å lage en revidert utgave (FPS-R), som skulle gjøre skalaen lettere å tolke og sammenligne med andre smertescoringsverktøy. Dette ble gjennomført av Carrie L. Hicks og hennes forskningsgruppe, og artikkelen ble gitt ut i år 2000 (4). Det ble gjennomført 3 studier for å revidere FPS til en skala med 6 ansikter, som skulle være mer praktisk å score, i forhold til numeriske skalaer. Ved hjelp av en datamaskinanimert versjon av FPS (SAFE –The Sydney Animated Facial Expressions) utviklet av Champion og kolleger (4) ble psykofysiologiske metoder brukt for å identifisere fire ansikter som representerte like intervaller mellom hvert ansikt, og verdiene som representerer minst og mest smerte (4).

FPS-R skalaen består i dag av 6 ansikter, hvor hvert ansikt representerer en økning i smerteintensitet fra venstre mot høyere. Skåringsanalysen for hvert ansikt er 0-2-4-6-8-10 når man oversetter ansiktene til numerisk skala. 0 er ingen smerte, og 10 er den sterkeste smerten man har opplevd; konferer figur 1 i (4). For å utvikle skalaen med 6 ansikter ble det gjennomført en serie av kryss-modalitet-kobling studie hvor de koblet tall og bilde, ved hjelp av personer som forstår å bruke begge deler (4). I den første studien ble det kun brukt én modalitet hvor voksne ble presentert for ansiktsskalaen. De skulle matche et tall opp mot hvert enkelt ansikt i skalaen. De valgte å benytte voksne i denne fasen, da oppgaven var ansett som for komplisert for barn. Alle deltakerne hadde god erfaring med barn mellom 5-12 år, og de ble informert om at de var med på å utvikle en skala som i fremtiden ville bli brukt for å måle smerte hos barn. I studie 2 og 3 (4) ble kryss-modalitet-kobling brukt på barn som skulle angi smerte i forhold til en ekte smertesituasjon. Hvert barn matchet sine verdier på ansiktsskalaen mot en numerisk skala som VAS, som er et godt dokumentert og pålitelig gyldig smerte-verktøy. Det ble funnet en sterk positiv korrelasjon ($r=0,93$, $p<0,001$) mellom smerteintensitet målt med VAS og med FPS-R i alle aldersgrupper (5-6 år, 7-8 år og 9-12 år). Den sterke positive korrelasjonen mellom smerteintensitet mellom VAS og FPS-R tyder på at FPS-R er et godt verktøy, og forskerne mener også at skalaen kan betraktes som en intervall-skala over et kontinuum, slik som VAS (4, 61).



Figur 1 (4).

Øverste rad viser FPS som scores 0-6. Nederste rad viser FPS-R som scores 0-2-4-6-8-10 eller 1-2-3-4-5. Se også www.iasp-pain.org/FPSR

6.2.2 Smerteverktøy basert på numeriske verdier og farger

Totalt var 11 artikler i litteratursøket som hovedsakelig tok for seg de numeriske skalaene (20, 21, 52, 66, 67, 69, 73, 76, 85, 86, 91). De fleste omhandlet mer enn én numerisk skala. Det var totalt 11 artikler som sammenlignet ulike numeriske skalaer med ansiktsskalaer (5, 7, 10, 15-17, 25, 26, 65, 68, 72). I avsnittene under gjennomgås fire vanlige numeriske skalaer. Sammenligning av ansiktsskalaer og numeriske skalaer tas opp i tilslutt i resultatdelen.

Smerteskalaer som er numeriske kan være verbale, ved at man spør barnet hvor ille smerten er på en skala fra 1-10. Numeriske skalaer kan også være skriftlige, ved at barnet får fremstilt og skal vise smerten på en skala for eksempel på papir. En forutsetning for å kunne bruke numeriske skalaer er at barnet klarer å rangere et tall mellom 0 (minst smerte) til 10 (verst smerte) (7). Slike skalaer er blitt validert for rapportering av akutt smerte i postoperative sammenhenger (68), ulykker, og akutte smertesituasjoner (46). Når man skal bruke disse skalaene er det viktig å huske at en andel av befolkningen har dyskalkuli, som gjør at de kan ha problemer med enkel numerisk tolking (7).

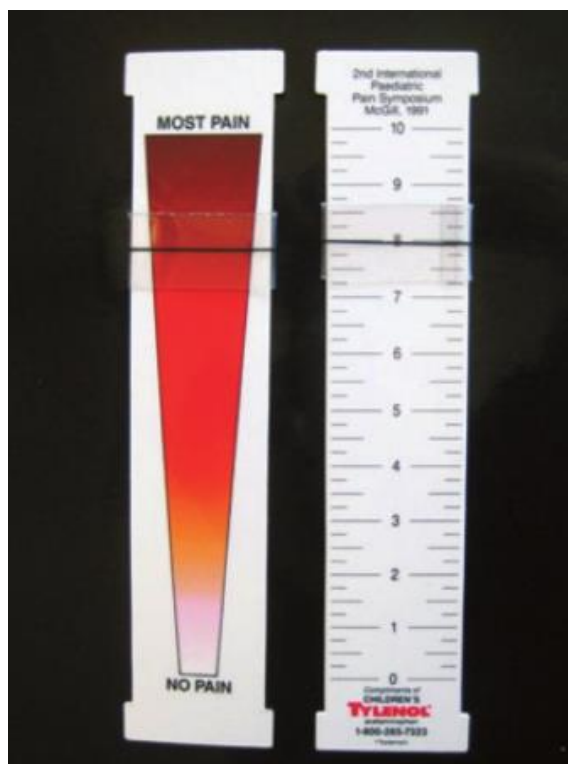
Det er spesielt to skalaer som baserer seg på tall. Disse går fra 0-5 eller 0-10. Numerical rating scale (NRS) (52, 67, 69, 76) og Visual analogue scale (VAS) (73, 76, 85, 86, 91).

Det er også laget to forenklete versjoner der tallinjen er visualisert ved hjelp av farger, Color Analogue Scale (CAS) (20, 52, 73) og Rainbow Pain Scale (RPS) (66, 80). Det er diskutert i hvor unge aldersgrupper disse verktøyene kan brukes. Studier har vist at CAS og RPS trolig kan brukes i yngre aldersgrupper, fra rundt 5 år (89), mens de kreves mer utviklet kognitive evner for å benytte NRS og VAS (89, 90).

Color Analogue Scale (CAS) er en type VAS-skala som har en kjegleformet fargegradering med rød farge på den ene siden, og en numerisk skala på den andre siden. På skalaen er det en bevegelig slider. Barnet får se den fargede siden av instrumentet, og skal bevege slideren til det punktet de mener representerer sin smerte (rødere farge indikerer sterkere smerte), og bevegelse vertikalt oppover på skalaen indikerer øktende smerteintensitet. Når barnet har markert hvor sterk smerten er, leser helsepersonell av på den andre numeriske siden av skalaen for å finne det korresponderende tallet. Skalaen er 10 cm lang, scoringen går mellom 0 og 10, med 0,25 cm differanse mellom hver linje (20, 72).

Fargeskalaen egner seg spesielt for barn 7 år og eldre (73), men studier har vist at den kan brukes fra 5 år og oppover.

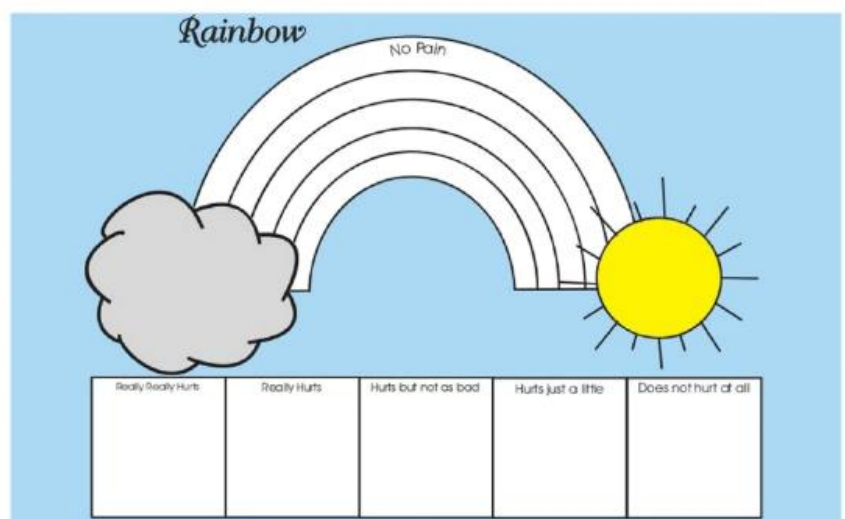
Blake Bulloch og medarbeidere (20) gjennomførte i 2009 en prospektivt studie hvor de fulgte 170 pediatriske pasienter i alderen 5-16 år i akuttmottaket. Hensikten med studien var å evaluere CAS-verktøyet for akutt smerte, generelt og fordelt på traumatiske og ikke-traumatiske tilstander. CAS skulle scores hvert 30 minutt, i 2 timer, totalt fire ganger eller til barnet ble skrevet ut fra akuttmottaket. For å evaluere repeterbarheten av en scoring ble det i tillegg gjort en måling 1 minutt etter den første vurderingen på tid 0 (akkurat i det barnet kom inn på akuttmottaket), og for hvert 30 minutt, også ved tid 31 minutter. De antok at det ikke ville være en endring i smerteintensitet i løpet av 1 minutt. Barnet fikk ikke se sine tidligere scoringer, for å unngå sammenligning av verdiene. Etter den første smertescoringen vurderte den behandlende legen behovet for smertestillende, før videre scoring ble gjennomført ved de aktuelle intervallene. Intraklasse korrelasjons-koeffisient (ICC) ble brukt for å evaluere repeterbarheten til 1-minutts intervallmålinger. Den gjennomsnittlige initiale CAS-scoringen for pasientene kombinert var 6,44 ($\pm 2,4$). Gjennomsnittlig score for barn med traumatiske skader var 5,94 ($\pm 2,6$) og for ikke traumatisk smerte 6,82 ($\pm 2,2$; $p=0,02$). 68% av pasientene ble skrevet ut etter 60 minutter.



Figur 1 (20). Color analog scale (CAS)

De selvrapporterte smertescorene var generelt reproducerbare ved 1 minutt fra initial presentasjon (ICC=0,97, 95% CI: 0,95-0,98) og etter 30 minutter (ICC=0,98, 95% CI: 0,97-0,99). Med traumatisk smerteetiologi var scorene reproducerbare ved initial presentasjon og 30 min oppfølging, ICC 0,98 (95% CI: 0,96-0,99) og 0,98 (95% CI: 0,97-0,99). For smerte med ikke-traumatisk etiologi var reproducerbarheten ved initial presentasjon og etter 30 minutter ICC 0,96 (95% CI: 0,91-0,98) og 0,98 (95% CI: 0,96-0,99). Resultatene viste at CAS egner seg godt i akutte situasjoner i aldersgruppen 5-16 år, og at skalaen har høy validitet både for traumatisk og ikke-traumatisk smerte (20).

Rainbow pain scale (RPS) er en skala basert på farger. Istedenfor å skulle bruke systemer som kan være for kognitivt utfordrende for små barn, ble fargeskalaen utviklet for at det skulle være enkelt for barn å rapportere smerte. Johnson, McQueen og Huetting (62), fant ut at selv 2-åringer har kunnskap om farger. De testet barn som enda



Appendix (66): Rainbow Pain Scale

ikke hadde utviklet språk for de ulike fargene. De fant ut at barna kan identifisere farger ved å peke på de, selv om de ikke klarer å fortelle hvilke farge det er. De mente derfor at barn kan identifisere farger, før de har evnen til å gi uttrykk for fargen verbalt. Rainbow Pain Scale (RPS) ble utviklet av Beverly Hillstrom (80), mor til Amanda som 2,5 år gammel fikk diagnostisert hjernestamme-gliom. Moren forklarte datter om svulsten ved å si at det var som hun hadde en «ball» i hodet som ikke skulle være der. Over de neste ukene beskrev Amanda «ballen» som at den hadde en farge hver dag. Det var totalt 5 farger i alt. Moren farget alle fargene på papir, og viste de til Amanda, og spurte om dette var fargene hennes. Gjennom resten av behandlingen klarte hun å uttrykke smertenivåer i farger. Det høyeste smertenivået beskrev hun som grønt, og det laveste smertenivået som rosa. Amanda døde i 1998, 3 år gammel. Sykepleierene på avdelingen spurte om de kunne utvikle en skala basert på denne, ettersom de så hvor mye den hadde hjulpet Amanda. Istedenfor å forvente at barn skulle bruke tall eller bilder for å rapportere smerte, mente de at det var lettere for barn å identifisere farger. Målet med RPS

var å gi hvert en enkelt barn et personlig scoringsformat for erfaring og uttrykk for smerte, ved at de kunne velge sine egne 5 farger som sitt uttrykk for en graderingsskala fra ingen til verst mulig smerte (66). Det langsikte målet er å gjennomføre psykometrisk testing av skalaen, siden det ikke tidligere har blitt rapportert en formell vurdering av verktøyet. Det ble gjennomført en studie av Paula Mahson og hennes forsikningsgruppe i 2015 (66). Studien ble gjennomført på en gruppe på 49 barn med kreftsmarter i aldersgruppen 5-10 år, og de sammenlignet RPS med FSP-R. Barna ble spurt fem spørsmål i henholdt til smerte, og ble bedt om å knytte en farge opp mot aktuell smerte, og deretter tegne en regnbue. Regnbuen trengte ikke å inneholde de opprinnelige fargene til regnbuen, men det skulle være farger barnet valgte selv. Barna graderte også smerte ved hjelp av FSP-R. Barna ble bedt å gjøre dette ved tre ulike anledninger de var på sykehuset. Korrelasjonen mellom RPS og FPS-R var 0,96 ved første visitt, 0,97 ved andre visitt og 0,93 ved tredje visitt. En krysstabell viste 92% enighet mellom valgene gjort av barna ved første visitt, 91% andre visitt og 87% tredje visitt. Det viste seg at noen av barna som oppga ingen smerte på RPS, anga ansikt nummer to på FPS-R, dette gjaldt spesielt på tredje visitt (66).

Numerical rating scale (NRS) er en ren tallskala som gir pasienten mulighet til å kvantifisere smerten fra 0–10 eller 0–5, der null representerer «ingen smerte» og fem eller ti representerer «verst mulige smerte». Litteraturen om bruk av NRS hos barn er hittil ikke dekkende. Det finnes ingen standard formulering eller pålitelig informasjon om den yngste aldersgruppen NRS er passende for. Det er likevel publisert studier som har sammenlignet NRS med VAS og/eller ansiktsskalaer. Resultatene fra disse studiene indikerer at barn i alderen 8 år eller eldre som kan forstå en numerisk rekkefølge meningsfullt kan bruke NRS (10, 52, 69, 76). Det er imidlertid andre studier som hevder at NRS ikke kan brukes før 9 års alder (16). Ytterligere data og nye studier er altså nødvendig for å vurdere om NRS-skalaen kan brukes av yngre barn.

NRS-skalaen trenger ikke å fremlegges visuelt til pasienten. Oftest brukes NRS i klinisk praksis ved at helsepersonell spør barnet hvor sterk smerte de har, på en skala fra 0–10 eller 0–5, altså Verbal-NRS (VNRS) (76). Bruk av NRS krever tall- og telleferdigheter og evne til å estimere mengder i numeriske termer, noe som krever betydelig kognitiv kapasitet. Fordelen med NRS-skalaen er at den er rask å bruke (52, 76).

Castarlenas og medarbeidere gjennomførte en studie som evaluerte bruk av den verbale utgaven av NRS hos 6–8 år gamle barn (52). Deres funn tydet på at skalaen egnet seg i denne aldersgruppen. Ingen av barna som deltok trengte spesifikk trening for å forstå hvordan de

skulle bruke skalaen. Validiteten til skalaen ble vurdert i fire ulike situasjoner som var den verste smerten barnet hadde hatt de siste 3 mnd, og tre tenkte spesifikke smertefulle hendelser. Studien evaluerte også validiteten av verbal NRS i forhold til tre andre skalaer (FPS-R, CAS og VAS), og så på hvilken skala barna foretrakk å bruke. Spredningsplott som er et statistisk diagram for å vise verdien fra to variabler i et dataset, i dette tilfellet (NRS vs. CAS), (NRS vs. FPS-R), (NRS vs. VAS) var lineært. Pearsons korrelasjon mellom NRS og de andre skalaene var sterk og positiv ($r=0,73-0,86$) samlet, noe som støtter *konvergent* validitet (at det er høyt samsvar mellom ulike metoder). Tallene sier ingenting om skalaene i seg selv er nøyaktige og om de kan skille mellom ulike smerteopplevelser. Man må derfor også teste skalene i en mer ekte klinisk setting, for å finne ut om skala kan brukes ved alle typer smerte (52).

Få studier har evaluert NRS' psykometriske egenskaper. En studie av Page (68) undersøkte NRS' *konvergente* validitet ved å sammenligne NRS med andre smerteskalaer som FPS-R og CAS. Studien undersøkte *diskriminerings*validitet ved å sammenligne NRS-skalaen med andre skalaer som måler lignende, men ikke identisk innhold, i dette tilfelle funksjonsnedsettelse. *Konstrukt*validiteten (om skalaen faktisk måler det den skal) ble undersøkt ved å bruke konvergent validitet, diskrimineringsvaliditet og sensitiviteten til testen over tid. Studien viste at NRS hadde god konvergentvaliditet med FPS-R ($p<0,01$). De fant signifikante, moderate korrelasjoner mellom smerteintensitet og funksjonsnedsettelse 2 uker etter det kirurgiske inngrepet ($p<0,001$) (69).

En studie av Mirò og Castarlenas (67) viste også korrelasjon mellom FPS-R og NRS for barn i alderen 8–12 år ($r=0,58$, $p<0,001$), også for postoperativ smerte. Testing av NRS' psykometriske egenskaper viste god konvergent, diskriminerende og kriterie-validitet. Gyldigheten av NRS brukt hos barn støttes også av von Baeyer og medarbeidere (76) som rapporterer flere eksempler fra barn som gjennomgår vaksinerings eller kirurgi. Samlet demonstrerer dette at NRS har en tilstrekkelig konvergent gyldighet pga. sine funn. De konkluderer med at NRS lett kan brukes i klinisk praksis for å vurdere smerteintensitet hos barn 8 år eller eldre.

Visual analogue scale (VAS) er en visuell skala, der smerteintensitet angis som en fysisk størrelse, ikke et verbalt utsagn (91). Skalaen fremstilles vanligvis som en horisontal 100 mm lang linje hvor pasienten blir bedt om å rangere smertens størrelse ved å plassere en markør mellom 0 og maksimum, som enten er 10 eller 100. Punktet hvor markøren settes skal

tilsvare nivået av smerteintensitet til barnet. 0 representerer «ingen smerte», mens 10 eller 100 representerer «verst tenkelige smerte».

Det er noe komplekst for barn å forstå VAS, og skalaen anbefales derfor for barn som er 8 år gamle eller eldre. Imidlertid hevder en artikkel av McGrath at skalaen kan brukes av barn over 5 år (11). Shields og medarbeidere (85) ønsket i en studie å undersøke gyldigheten til VAS for barn i alderen 5–14 år, fordi tidligere artikler de hadde gått igjennom mente at 13% av barna ikke kunne forstå skalaen, og at 39 % svarte feil på skalaen (85). De undersøkte også om måten man lærer bort bruken av VAS-verktøyet har noe å si for om barna klarer å bruke verktøyet. Assosiasjonen mellom barnets alder og evnen til å forstå konseptet av VAS ble også evaluert.

106 barn med kutt som krevde suturering etter lidokaininjeksjon som lokalbedøvelse deltok i studien. To verktøy ble brukt for å vurdere baseline smerte og lidokain-injeksjonssmerte: en 5-punkts Likert-skala som er en fem-punkt skala hvor barnets smerte beskrives som «ingen smerte», «litt smerte», «moderat smerte» «mye smerte» eller «verst tenkelig smerte» og VAS 100 mm horisontal skala.. Forsøkspersonene ble bedt om å markere på VAS-skalaen og Likert-skalaen ved to tidspunkt: forventningsfasen, dvs. før prosedyren ble utført, og injiseringsfasen dvs. mens lidokain ble injisert. Barna fikk også i oppgave å markere på linjen hvor vondt selve skademekanismen hadde gjort, og de fikk en del andre spørsmål knyttet til smerte og tidligere smerteopplevelser. Resultatene viste at måten man lærte barna å bruke VAS-verktøyet ikke hadde innvirkning på hvor mange av dem som markerte svarene sine «riktig» på skalaen sammenlignet med Likert-skalaen. Undervisningsmetoden hadde heller ikke effekt på hvor mange barn som kunne forstå skalaen. Forskerne fant ut at barna som var i stand til å forstå VAS var eldre (gjennomsnitt = 9,8 år, SD= 2,8) enn de som ikke gjorde det (gjennomsnitt = 8,2 år, SD= 2,5 år). De fleste av barna brukte kun endepunkt og midtpunkt av skalaen når de skulle rapportere smerte, og noen av barna satte markøren på samme punkt hver gang, Totalt var det bare ca. 1/3 av barna som var i stand til å merke «riktig» VAS-verdi i forhold til smerten de anga på Likert skalaen, altså å forstå konseptet med VAS. Når de sammenlignet verdiene i forventningsfasen og injeksjonsfasen fant de at Likert- og VAS-scorene samsvarte i 56,5% av tilfellene (eks Likert-score rapportert av barnet i injeksjonsfasen var høyere enn i forventningsfasen og VAS-score rapportert av det samme barnet var også høyere i injeksjonsfasen enn i forventningsfasen)(85). VAS er altså en ganske kompleks

skala som krever at brukeren kan oversette en subjektiv sensorisk erfaring til et lineært format (86). Forskerne konkluderte med at VAS mulig ikke er et valid verktøy for yngre barn, og at det kan være en fordel å benytte måleverktøy mer tilpasset yngre aldersgrupper (85).

Shields og medarbeidere ønsket også å undersøke hvordan kognitive evner og alder påvirker evnen til å bruke VAS (86). 40 barn i alderen 5,0 – 6,8 år med ulik bagrunn (72% kaukasisk, 20% afro-amerikansk og 8% andre ikke-latinamerikanere) ble inkludert. Sosioøkonomisk status hos barnas foreldre ble estimert, barna gjennomgikk ulike IQ-tester, og barnas skår ble skalert etter en standardisert tabell. Barna ble så presentert for et hefte med 14 ulike sirkler i tilfeldig rekkefølge, og ble bedt om å rangere størrelsen på hver sirkel på et medfølgende VAS-verktøy. Barna ble forklart og vist at den minste sirkelen representerte venstre endepunkt på VAS og at den største sirkelen representerte høyre endepunkt på VAS skalaen. Lineær regresjonsanalyse ble brukt for å bestemme om det var en lineær sammenheng mellom hver individs respons (opplevd sirkelstørrelse) og faktisk sirkelstørrelse.

Resultatene viste at 42% (n=17) klarte å bruke skalaen riktig. De som ikke klarte å bruke skalaen riktig brukte kun endepunktene eller midtpunktet, markerte på omtrent samme sted for hver sirkel, men varierte med den vertikale høyden for å tilsvare endring i sirkelstørrelse. Evnen til å utføre serieringsoppgaven var signifikant assosiert med barnets evne til å bruke VAS ($p=0,001$). De fant at alder $>5,6$ år koblet sammen med en estimert IQ > 99 var den beste prediktoren for barnets evne til bruke VAS (86).

6.2.3 Ansiktsskalaer versus numeriske skalaer

Man ville forvente at det var forskjell i effekten av ansiktsskalaer, der emosjonelle uttrykk symboliserer smerte, og numeriske skalaer, som krever at barnet «oversetter» smerten sin til et tall, en fargeintensitet, eller en fysisk lengde.

En studie gjort av Tsze og hans forskningsgruppe (72) i 2013 sammenlignet FPS-R og CAS for akutte smerter hos barn med gjennomsnittalder 9,2 år (SD 3,8 år). 50,5% av barna var <8 år. Forskerne så på validitet og pålitelighet mellom ansiktsskalaen og fargeskalaen i henhold til barnas kjønn, alder og etnisitet. De hadde en frisk kontrollgruppe (uten smerter) som hadde samme alder, kjønn og etnisitet. Barna med smerter skulle når de ankom akuttmottaket angi smerten sin både på FPS-R og CAS. Det ble gitt en dose smertestillende medikament, og testingen ble gjentatt 30 minutter etter dette. Undersøkelsen ble gjort på nøyaktig samme måte for kontrollgruppen, men disse fikk ikke smertestillende. Kontrollgruppen ble brukt for

å undersøke om testen faktisk måler smerte. Det ble forventet at scoringsverdien skulle gå ned hos barnegruppen med smerter etter de hadde fått smertestillende, mens man forventet ingen forandring i verdiene til kontrollgruppen.

Studien tok altså for seg gyldigheten av FPS-R og CAS, i form av *konvergent* validitet (hvilken grad to ulike skaler som er ment å måle den samme tingen, smerte, gir lignende resultater), *diskriminerende* validitet (om skalaen måler den parameteren den er ment å måle, og ikke andre ting som for eksempel angst), og *respons* (i hvilken grad en test oppfører seg på en måte som samsvarer med den forventede). De vurderte også påliteligheten til begge smerteskalaene.

De 620 pasientene som var med i studien ble delt inn gruppene yngre (4–7 år) versus eldre (8–17 år), jenter versus gutter, og latinamerikansk versus ikke-latinamerikansk etnisitet. Det ble også inkludert subgrupper, hvor man undersøkte alle typer validitet. Disse subgruppene var: 4, 5, 6, og 7-åringer, afroamerikansk versus «hvit» etnisitet, og engelsk versus spansk primært talespråk. Forfatterene vurderte at skalaene korrelerte med hverandre, dersom forskjellene i scoringsverdi for CAS og FPS-R var <3 poeng for mer enn 80% av barna. For å undersøke *diskriminerings*-validitet sammenlignet de initial gjennomsnittscore av FPS-R og CAS hos barn med smertefulle tilstander sammenlignet med de uten smerter. Analysen ble gjort i alle aldersgrupper. *Responsiviteten* ble bestemt ved å sammenligne initiale gjennomsnitt av FPS- og CAS-score med deres respektive postanalgesi smertescore. Kontrollgruppen ble vurdert på samme måte, men med en forventning om at det ikke skulle være noen forskjell mellom den første målingen og etter 30 min. Korrelasjonskoeffisientene mellom FPS-R og CAS var generelt høy. Korrelasjonen var høyere i gruppen med eldre barn ($p < 0,0001$), hos jenter sammenlignet med gutter ($p < 0,01$), og hos «hvite» sammenlignet med afroamerikanske ($p < 0,0001$). Korrelasjonen mellom skalaene var 0,82 hos 6 åringer, men 0,73 hos 5 åringer. Verdiene var altså mindre sikre for de minste barna. *Diskriminerings*-validiteten var god, dvs. at scoringsverdiene for både FPS-R og CAS var høyere ($p < 0,0001$) i barnegruppen som hadde smerter. Det var lignende differanse for alle subgruppene.

Responsiviteten viste at barna med smerter rapporterte lavere verdier for både FPS-R og CAS etter de hadde fått smertestillende, mens i kontrollgruppen var det ingen endring (72). De to smerteverktøyene målte altså faktisk smerter, og korrelerte akseptabelt med hverandre.

Robin Luffy og Susan K Grove (65) undersøkte Oucher pain scale, FSP-R og VAS mht. *samtidig validitet* (concurrent validity), som sier noe om hvor godt resultatene samsvarer mellom ulike måleverktøy, og *pålitelighet* (reliability) som beskriver målesikkerhet når

samme måling gjentas mange ganger med samme premisser (72). De undersøkte også hvilken skala barna foretrakk å bruke. Inkludert i studien var 100 afroamerikanske barn i alderen 3–18 år. Resultatene viste at i aldersgruppen 3–7 år foretrakk 63% av barna å bruke FSP-R, 28% foretrakk Oucher-skalen, og 9% foretrakk VAS. Validiteten av skalene var for FSP-R 57%, Oucher 50%, og VAS 33%. Påliteligheten til skalaene var for FSP-R 46%, Oucher 57%, og VAS 28%. I alderen 8–12 år foretrakk 47% av barna å bruke FSP-R, 28% Oucher, og 25% VAS. Validiteten til skalaene hos disse eldre barna var for FSP-R 81%, Oucher 83%, og VAS 81%. Påliteligheten til skalaene var for FSP-R 78%, Oucher 81%, og VAS 58%. For barn i skolealder var det ingen signifikant forskjell mellom de tre ulike verktøyene. Generelt var validiteten størst for FSP-R, etterfulgt av Oucher, og validiteten og påliteligheten økte med økende alder hos barna for alle verktøyene. I aldersgruppen 3–7 år var skalaen med høyest pålitelighet ansiktsskalaen Oucher, men FSP-R hadde høyest validitet (65).

Det har også blitt utviklet en «word descriptor scale» som har blitt beskrevet i «*Whaley and Wong (1987)*»(63) som er en VAS lignende skala. Skalaen består av 5 ord som brukes for å beskrive smerte («*no pain*», «*little pain*», «*medium pain*», «*large pain*», «*worst pain*»). Hvert ord er knyttet til et tall. Tallet er forbundet med de samme numrene som ansiktsskalaene. Om ordet representerer tallverdien det hører sammen med krever videre forskning (63). Juanita Fogel Keck (63) og hennes forskningsgruppe gjennomførte en studie hvor de sammenlignet ansiktsskalaer med ordbeskrivende skala hos 110 barn i tre ulike aldersgrupper (3-7 år , 8-12 år og 13-18 år). De fant ingen signifikant forskjell mellom de to skalaene i noen av aldersgruppene. Likevel viste seg at når de spurte hvilket verktøy barna likte best å anvende, var det ansiktsskalaen som kom best ut (83,3% i gruppen 3-7 år, 63,4% i gruppen 8-12 år).

Hovedgrunnlaget for å bruke ansiktsskaler fremfor numeriske skalaer hos barn er at ansiktsskalaer ikke krever størrelsesestimat eller seriering; dette er kognitive ferdigheter som de fleste barn ikke har før de går på skolen. Små barn mangler også erfaring med smerte, det er derfor mer individuelt hva barn tenker på som «verst tenkelig smerte». Noen barn kan tenke på smertetilstander de har hørt om, mens andre barn tenker på den spesielle smerten de nettopp har opplevd eller husker best. Dette problemet blir mindre når man bruker skalaene repeterte ganger i ett enkelt individ, og bruker den samme skalaen konsekvent (16).

6.3. Observasjonelle eller objektive smertescoringsverktøy

Objektive smerteskalaer benyttes når selvrapporteringsverktøy ikke er tilstrekkelig, eller som supplement til selvrapporteringsverktøy. Disse skalaene er hovedsakelig beregnet for barn <4 år (54), men andre mener de bør benyttes for barn <6 år (12). De kan også brukes hos eldre barn (64). I søket var det totalt 18 artikler som tok for seg observasjonsverktøy (2, 3, 12, 14, 18, 22, 27-29, 45, 54, 56, 60, 64, 75, 83, 84, 88). Fem artikler sammenlignet observasjonsverktøy med selvrapporteringsverktøy (13, 34, 57, 78, 87).

Ansiktsuttrykk har vært den variabelen som er mest forsket på med utgangspunkt i smerte. De vanligste endringene er spenning i ansiktsmuskulaturen, sammenkniping/lukking av øyne, utviding av nesebor, bøyning av øyebryn og stramming av tungen (2). Gråt hos barn har en viktig rolle i å varsle voksne om de minste barnas behov. Det er ulike karakteristika ved gråten som for eksempel frekvens på gråten, hvor mye energi barnet gråter med, melodien på gråten, og anstrengthet. Det er likevel store vanskeligheter med diskriminere for eksempel smerte og angst basert på gråt. Dette har til nå ikke vært mulig, og gråt må derfor ses i sammenheng med andre parametere (2, 27). I en studie av Dubois og kolleger (55) hvor de undersøkte utviklingen av vokalt og verbalt uttrykk for postoperativ smerte viste resultatene at det vokale uttrykket utvikler seg i henhold til barnets kronologiske alder. Barn under 24 mnd brukte kun gråt som et uttrykk for smerte, mens etter hvert som barna ble større ble smerten uttrykt mindre i gråt og mer verbalt i form av ord (55). Observasjon av kroppsbevegelse har også blitt sett på som en parameter ved smerte, da det er observert at barn kan respondere med spasmer, rigiditet og muskelspenninger i ekstremiteter. Denne parameteren blir mindre tydelig når barnet blir eldre, og får bedre kontroll på ekstremitetene (2). Utrøstelighet og sovemønster har man sett hos barn med mer langvarige smertetilstander (2). De fleste av skalaene innen observasjonskategorien tar for seg disse parameterene i større eller mindre grad. Likevel er det diskutert om observasjonsverktøy er mindre egnet for langvarige og kroniske smertetilstander, fordi de er sårbare for endring i atferdsmønsteret. Barn som er utsatt for kronisk smerte vil etter hvert kunne ha en tendens til å tilvenne seg smerten og dermed ikke vise de samme tegnene med tiden. Det er dermed vanskelig å observere påliteligheten til disse variablene i tilfeller med kroniske smerter (17).

FLACC (*figur 1 (64)*): Ble først utviklet i 1997 og brukes på barn fra 2 mnd til 7 år (18). Det er i dag et av de mest brukte observasjonsverktøyene. Formålet med skalaen var at man skulle ha en måte og raskt kunne evaluere postoperativ smerte (64). FLACC har blitt anbefalt

som førstevalg for postoperativt smerte på sykehus (27). Det har blitt gjennomført yttligere psykometrisk testing, og FLACC har vist seg å være gyldig, pålitelig og mulig å bruke i en rekke andre situasjoner som for eksempel traume (64), og kortvarige akutte smerter for eksempel ved vaksinasjon av barn mellom 18-22 måneder (60). FLACC tar for seg de ulike atferdsparametere gråt («cry»), grimaser («facial»), aktivitet («activity»), bevegelse av barnets ben («legs») og grad av trøstbarhet («consolability»). Hver parameter scores mellom 0-2, som resulterer i en totalscore mellom 0-10 (18, 60). Skalaen har en cut-off verdi på 3 (27). FLACC ble testet for akutt smerte av blant annet Kochman (64) ved en prospektiv observasjonsstudie, og de undersøkte interrater-reliability¹ ved at to sykepleiere evaluerte FLACC på samme pasient samtidig. De undersøkte også FLACC-score like før, og 30- og 60 minutter etter administrasjon av smertestillende. De fant at det var betydelig til nesten perfekt samstemthet mellom sykepleierne, 90% (27/30) for de individuelle komponentene i skalaen (rangert fra 0,63-1,00) og 100% for totalscorene (0,85-0,96) mellom de ulike sykepleierens scoringsverdi. Det var også en betydelig reduksjon i poengscore etter administrasjon av smertestillende etter 30 min (2,00 95% CI: 1,61-2,39) og 60 minutter (1,14 95% CI: 0,79-1,50). Denne forskjellen var signifikant med $p < 0,0001$ for alle sammenligninger (64).

Lignende funn ble også gjort av Sandra I. Merkel og Terri Voepel-Lewis (82) som gjennomførte en studie med 89 barn i alderen 2 mnd til 7 år. Interrater-reliabiliteten og validiteten ble undersøkt ved å gi analgetika og undersøke FLACC etter 10- (iv), 30- og 60 minutter (iv og p.o) og ved å sammenligne FLACC-score med OPS-score. Det var høy korrelasjon mellom to ulike observatørers målinger $r=0,94$ $p < 0,001$, og dermed god interrater-reliabilitet. Dette gjaldt også for de ulike kategoriene hver for seg. Prealgesi FLACC scorerer ($7,0 \pm 2,9$) var signifikant høyere enn postalgesi verdier ved 10 min ($1,7 \pm 2,2$), 30 min ($1,0 \pm 1,9$) og 60 min ($0,2 \pm 0,5$) ($p < 0,001$ for hvert tidsintervall). Det ble ikke funnet signifikant forskjell mellom de ulike tidsintervallene 10, 30 og 60 minutter. OPS og FLACC score hadde en signifikant positiv korrelasjon; dette støtter at begge skalaene måler smerte ($r=80$ $p < 0,001$) (82).

Voepel-Lewis og kolleger (75) undersøkte validiteten og reliabiliteten til FLACC brukt postoperativt for barn med kognitiv svikt. De undersøkte 79 barn i alderen 4-18 år som skulle gjennom generell- eller ortopedisk kirurgi. Etter kirurgi, men før de fikk smertestillende, ble barna filmet, og de gjennomførte 140 observasjonsfilmer. Filmen ble senere gjennomgått av

¹ Inter-rater reliabilitet: Uavhengig av hverandre får ulike personer samme resultat ved bruk av måleverktøyet.

den behandlende sykepleieren og en forelder, som vurderte barnets smerte samtidig ved å bruke FLACC og en VAS. De barna som fikk smertestillende ble filmet og scoret en gang til. To sykepleiere, blindet for om smertestillende var gitt, så videoene i tilfeldig rekkefølge og brukte FLACC. 50 videoer ble gjennomgått på nytt 2 måneder senere for å dokumentere intrarater-reliabilitet. Det ble funnet at FLACC hadde god interrater-reliabilitet, og god konstruksjonsvaliditet ble påvist gjennom en reduksjon i FLACC etter smertestillende. Samtidig validitet ble støttet av en sterk korrelasjon mellom FLACC og foreldrenes VAS-score. Likevel viste det seg at foreldrenes score oftere var høyere enn FLACC scorene som ble gjort av behandlende sykepleier og av sykepleierne som var blindet for om smertestillende var gitt. Sykepleierne som var blindet ga en lavere score enn behandlende sykepleier. Både behandlende sykepleier og sykepleierne blindet for smertestillende ga en lavere FLACC-score etter administrering av smertestillende ($p < 0,001$). Alle barna ble også testet for evnen til å selvrapportere smerte. 16 barn ble vurdert at de klarte selvrapportering, men bare 8 stykker gjennomførte. Forskerne fikk derfor ikke undersøkt sammenhengen mellom FLACC og selvrapportering (75). FLACC-skalaen har også blitt validert i forhold til oversettelse til andre språk, blant annet kinesisk (45).

Categories	Scoring		
	0	1	2
Face	No particular expression or smile	Occasional grimace or frown; withdrawn, disinterested	Frequent to constant frown, quivering chin, clenched jaw
Legs	Normal position or relaxed	Uneasy, restless, tense	Kicking or legs drawn up
Activity	Lying quietly, normal position, moves easily	Squirming, shifting back and forth, tense	Arched, rigid or jerking
Cry	No cry (awake or asleep)	Moans or whimpers; occasional complaint	Crying, steadily, screams or sobs, frequent complaints
Consolability	Content, relaxed	Reassured by occasional touching, hugging, or being talked to; distractible	Difficult to console or comfort

FLACC smertescoringsverktøy. Omarbeidet Figur 1 fra kilde (64)

Children's Hospital of Eastern Ontario Pain Scale (CHEOPS): Ble utviklet i 1985 av McGrath et al (22). Skala ble utviklet for å evaluere postoperativ smerte hos barn mellom 1-7 år. Elementene i skalaen er basert på beskrivelser av de mest observerte atferdene hos barn med smerte, som er beskrevet av erfarne postoperative sykepleiere. Scoringsverdiene strekker seg fra 4 (ingen smerte) til 13 (verst smerte), og består av 6 atferdskategorier: «cry», «facial», «verbal», «torso», «touch» og «legs». Hver kategori evalueres i tre eller fire grader (ingen smerte, mild, moderat og alvorlig smerte) (54). En CHEOPS score rundt 7 poeng indikerer i de fleste studier at barnet trenger smertebehandling (22). CHEOPS scores ved at observatøren

observerer barnet 5 sekunder for hver kategori, etterfulgt av 25 sekunder for å registrere total poengsum. CHEOPS ble vist å være en pålitelig skala, med gjennomsnittlig interrater-korrelasjon fra 0,9-0,99. Skalaen viste god intern korrelasjon mellom de ulike variablene og total poengsum (54). Konstruksjonsvaliditet ble testet ved å vise at CHEOPS-scoren ble redusert etter administrering av smertestillende. Sykepleierne som gjennomførte scoren var de samme som ga smertestillende, slik at dette kan mulig ha påvirket resultatet. CHEOPS har også blitt sammenlignet med andre smerteskalaer som Objectiv pain scale (OPS) (6 mnd-3 år), og FPS og VAS (6,5-12 år). Pearssons-korrelasjon viste en verdi $r=0,743-0,921$ mellom de forskjellige skalaene. Forskerne anså derfor alle skalaene som gyldige for postoperativ smertevurdering av barn i alderen 6 mnd-12 år (54).

Suraseranivongse og medarbeidere (87) gjennomførte en kryssvalideringsstudie for å evaluere CHEOPS, FLACC, TPPPS og OPS hos 167 barn i alderen 1-5 år. De tok video av alle barna før operasjon og etter operasjonen, før analgesi ble gitt. Fire observatører så på videoene og gjorde en smerteevaluering ved hjelp av de fire skalaene. Sammenligningen av smertescorene før og etter kirurgi viste at alle skalaene hadde kapasitet til å diskriminere ($p<0,001$) mellom ulike grader av smerte. Samtidig validitet (concurrent validity) som sier noe om hvor godt resultatene mellom de ulike skalaene samsvarer viste moderat til god korrelasjon ($r=0,62-0,77$ $p<0,0001$). Det høyeste nivået av enighet mellom skala og den kliniske diskusjonen om de skulle behandle smerten ble vist med CHEOPS. Inter- og intrarater reliabilitet ble også testet og var høy for alle skalaene. Forfatterne anbefaler CHEOPS for smertevurdering i den umiddelbare postoperative fasen for barn i alderen 1-5 år (87). Beyer og kolleger fant at CHEOPS-score var generelt svært lave og ikke korrelerte godt med selvrappoteringsverktøy for barn i alderen 3-7 år. De fant også at CHEOPS score generelt var lave noen timer etter kirurgi (78).

Children's and infant's post operative pain scale (CHIPPS): er en skala som ble utviklet av Büttner og Finke (28). Den ble utviklet ved bruk av atferdsindikatorer for smerte hentet fra litteratur på området. Det er en skala som består av 5 atferdsvariabler (crying, facial expression, trunk, leg posture and motor restlessness). Hver variabel scores fra 0-2. Totalscoren er mellom 0 og 10, med en cut-off på 3. Dersom barnet har en score på 3 eller over bør barnet få smertebehandling (27, 28). Cronbach's alfa koeffisient for indre konsistens var 0,92 for småbarn (12-36 mnd). Interrater reliabilitetskoeffisienten var 0,93. Sensitivitet og spesifisitet ble regnet ut til å være 92-96% og 74-95%. Barn som ikke hadde smerter hadde en CHIPPS

score på mindre enn fire poeng. Skalaen er også praktisk ved at den kan gjennomføres på 15 sekunder (28). Skalaen er ment for barn i alderen nyfødt til 5 år, og for akutte prosedyresmerter (18). CHIPPS ble opprinnelig laget på engelsk, men har også blant annet blitt oversatt til portugisisk (28).

The Toddler-Preschooler Postoperative Pain Scale (TPPPS): Er en oppførsel/ observasjonsskala validert for bruk i postoperative settinger og smertevurdering av barn i alderen 1-5 år (83, 88). TPPPS består av 7 elementer delt inn i tre hovedkategorier: 1) Vocal pain expression, 2) Facial pain expression og 3) Bodily pain expression. Poengsummen er mellom 0-7 (18, 83).

Den første valideringen ble utført av Tarbell og kolleger (88) på 74 barn i alderen 1-5 år, som våknet opp fra anestesi etter kirurgi av hernie eller hydrocele. Barna ble observert i seks 5-min intervaller etter oppvåkning av anestesi. Behandlende sykepleier gjennomførte en smertevurdering ved hjelp av TPPPS-score. En annen, uavhengig sykepleier observerte 28 av barna for å dokumentere reliabilitet og interrater-reliabilitet, som ble dokumentert som god både mellom de ulike variablene og verktøyet som helhet. Indre konsistens ble bestemt ved å sammenligne de seks intervallpoengene for hver pasient, og ved å sammenligne totalscoren mellom pasientene. Det var stor korrelasjon mellom de individuelle og totalscorene, men korrelasjonen mellom de individuelle variablene varierte fra dårlig til god. Dette kan reflektere ulike smertekvaliteter som de ulike variablene måler. Konstruksjonsvaliditeten ble vist gjennom et signifikant fall i TPPPS-score etter postoperativ smertestillende som ble gitt til 25 barn. Studien demonstrerte foreløpig validitet og reliabilitet for TPPPS, men hadde noen svakheter. Blant annet besto studiepopulasjonen for det meste av kaukasiske gutter, og sykepleierne var ikke «blindet» for om smertestillende ble gitt. Forfatterne bemerket at selv om TPPPS ble funnet å være følsom for administrasjon av smertestillende var de usikre på om denne følsomheten skyldtes at barnet hadde mindre vondt eller om det var en følge av den beroligende effekten noen typer smertestillende har (88).

Objective Pain Scale (OPS): Er en skala som består av 5 variabler, hvor hver variabel scores mellom 0-2. Variablene er: Blood pressure, crying, movement, agitation, verbal or body expression. Skalaen inneholder ikke noen variabel for ansiktsuttrykk. Skalaen er ment å måle postoperative smerte hos barn i alderen 18 mnd til 10 år (27). Skalaen har ofte blitt brukt uten blodtrykksvariabelen. Totalscore er fra 0 til 8 eller 10, med en cut-off-verdi på 2 eller 3

avhengig om man bruker 4 eller 5 variabler (27). OPS ligner på CHEOPS, men inkorporerer kun 4 variabler og krever også dokumentasjon rundt prosentforandring i blodtrykk i forhold til pasientens «baseline». Skalaen har fått rapportert god interrater-reliabilitet og validitet, og den har også god korrelasjon med selvrappoteringsverktøy for barn i alderen 5-13 år (82).

Preverbal, Early Verbal Pediatric Pain Scale (PEPPS), original og Modifisert utgave

(M-PEPPS): Den opprinnelige skalaen PEEPS ble utviklet som et verktøy for å måle smerte hos barn som ikke har evnen til å selvrappotere lokalisasjonen og graden av postoperativ smerte (3, 84). Skalaen tar for seg syv kategorier som er ment å være representative når barnet har smerter *heart rate, cry, facial, consolability/state of restfulness, body posture, sociability, suckling/feeding*, og hver av de inneholder to til fire indikatorer som indikerer varierende responser og nivåer av smertestimuli (84) Beskrivelsene reflekterer atferd fra mild til alvorlig. Elementene i skalaen ble utviklet ved gjennomgang av litteratur på området, observasjon av postoperativ smerteatferd hos små barn, og diskusjon med en rekke leger og sykepleiere som arbeider med denne pasientpopulasjonen. I 6 av de 7 kategoriene kan scoringsverdiene variere fra 0-4. For kategorien suckling/feeding varierer scoringen fra 0-2. Totalscoren bestemmes ved å summere de kategoriske scoringsverdiene, og den totale scoren kan være mellom 0 og 26 (84). Interrater-korrelasjoner varierte fra 0,90 til 0,96, mens intrarater-korrelasjonene som måler gjentakelsesgyldigheten til en person som gjør samme test, varierte fra 0,96-0,98. Indre konsistens for totalskalaen som måler samsvar mellom ulike testledd og om disse kan representeres som et sett i en gruppe, og sier noe om skalaen reliabilitet (stabiliteten til skalaen) var 0,89. Interitem-korrelasjonen var rangert fra 0,15 for suckling/feeding og heart rate, til 0,88 for consolability og cry. I totalkorrelasjonen utgjorde suckling/feeding og heart rate 0,37 og 0,38. Å eliminere disse elementene vil bedre den interne konsistensen. Konstruksjonsvaliditeten til PEPPS (om verktøyet måler det som er ment å måle) ble vurdert ved å bruke en sammenkoblet t-test for å sammenligne smerteverdier før og etter administrasjon av smertestillende medisin (Tylenol). Det var en statistisk signifikant forskjell i pre- og postmedisinsk score $t=14,58$ $p<0,001$. Det ble rapportert at tallene støtter at PEEPS er et valid og pålitelig verktøy for å måle postoperativ smerte (84) . Siden scoringsverktøyet opprinnelig ble laget for å måle postoperativ smerte, inkluderer den mange variabler, og er tidkrevende å bruke. Det har derfor blitt utviklet en modifisert utgave som skal gjøre bruken av skalaen lettere i akutte situasjoner (3). Siden mange av akutt-pasientene ikke kan spise eller drikke ble denne kategorien ekskludert i den nye versjonen. Mange av pasientene som kommer inn på akuttavdelingen er også febrile og engstelige. Dette

kan påvirke hjertefrekvensen, og denne variabelen ble derfor også ekskludert i den nye modifiserte utgaven (M-PEPPS). I den nye utgaven kan derfor den totale scoringen variere fra 0-20 (3). Det har blitt gjennomført en analyse av Strout og hans forskningsgruppe (3) for å se på om den nye versjonen M-PEPPS er like pålitelig og troverdig for akutte smerter som den gamle PEPPS er for postoperativ smerte. Studien ble gjennomført på 118 barn i alderen 12-84 mnd. Inter-item korrelasjoner mellom de ulike variablene varierte fra totalscoren 0,850 for facial til 0,917 som totalscore for posture. Cronbachs alfa- koeffisienten til den nye skalaen ble funnet til 0,954. Retest påliteligheten som er instrumentets evne til å gi konsistente resultater over tid, ble ikke vurdert da de mente at smerteintensiteten forventes å variere, snarere enn å være stabil over tid. Det ble gjort en analyse for alle de ulike variablene, og dataene viste at alle parameterne som brukes er nyttige når man skal evaluere smerte hos barn, og at verktøyet måler det som er ment å måle, nemlig smerte (3).

Parameter	0	1	2	3	4
Facial	Relaxed facial expression		Grimace, brows drawn together; eyes partially closed, squinting		Severe grimace; brows lowered, tightly drawn together; eyes tightly closed
Cry (audible/visual)	No cry	Whimpering, groaning	Intermittent crying	Sustained crying	Screaming
Consolability/state of restfulness	Pleasant, well integrated	Distractible, easy to console, intermittent fussiness	Able to console, distract with difficulty, intermittent restlessness, irritability		Unable to console, restlessness, sustained movement
Body Posture	Body at rest, relaxed positioning	Clenched fists, curled toes and/or reaching for, touching wound or area	Localization with extension or flexion or stiff and non-moving	Intermittent or sustained movement with or without rigidity	Sustained arching, flailing, thrashing and/or kicking
Sociability	Responds to voice and/or touch, makes eye contact and/or smiles, easy to obtain or maintain; sleeping		With effort, responds to voice and/or touch, makes eye contact but difficulty to obtain or maintain		Absent eye contact, no response to voice and/or touch

Utarbeidet tabell Modified Preverbal, Early Verbal Pediatric Pain Scale (M-PEPPS) fra kilde (3)

Behavioral pain scale (BOPS) (*figur 1 (22)*): er en smerteskala som ble utviklet i 1996 ved Neurointensive care unit i Lund i Sverige (22). Skalaen er en blanding av Princess Margaret Hospital Pain Assessment Tool og Children's Hospital of Eastern Ontario Pain Scale (CHEOPS). Tidligere erfaringer hadde vist at mange observasjonsverktøy var vanskelig å bruke på grunn av verktøyets kompleksitet, at det var u håndterlig eller at det tok lang tid å score pasientene. I tillegg kunne det være vanskelig å tolke resultatene. Formålet med skalaen var å redusere antall indikatorer for smerte og forenkle scoringssystemet. Scoringsverdien regnes ut basert på barnets ansiktsuttrykk, vokalisering og kroppsbevegelser. Disse tre faktorene ble betraktet som de mest pålitelige for å tolke smerte, da forskerne mente at tidligere studier har vist en sterk sammenheng mellom disse 3 kategoriene og smerteevaluering. Scoringsverdiene for verktøyet er 0, 1, og 2. Summen av variablene vil være mellom 0 og 6 poeng (22). Pålitelighet og validitet er viktige kriterier for å vurdere et kvantitativt instrument. I observasjonssmertevurdering av barn er interrater-reliabilitet viktig, da dette gir pålitelige målinger uavhengig barnets alder og kjønn i tillegg til hvilken kliniker som utfører testen. Validiteten gjenspeiler at instrumentet måler barnets smerte slik at en endringer i smerteverdi reflekterer meningsfulle forskjeller i smerteopplevelsen.

Hesselgard og medarbeidere (22) undersøkte om BOPS var et troverdig og pålitelig instrument for måling av postoperativ smerte hos 76 barn som var mellom 1-7 år. BOPS ble i studien sammenlignet med CHEOPS. For å utforske om ulike sykepleiere vil få samme observasjoner uavhengig av hverandre (inter-rater reliabilitet) ville to trente observatører observere samme pasient samtidig, og sette en BOPS score individuelt. For å styrke resultatet skulle ulike personer gjennomføre evalueringen, og totalt var det 24 sykepleiere som utførte undersøkelsen. Sykepleierne delte ikke scoringsverdien de fikk. For å undersøke om verktøyet faktisk vurdere smerte, og ikke noe annet, ble barna observert før og tre ganger etter smertestillende medisin. Sykepleierne skulle selv vurdere, og ta en avgjørelse på når smertestillende skulle gis. BOPS ble evaluert rett før smertestillende ble gitt og 15-, 30- og 60 minutter etter inntatt smertestillende. Det ble gjennomført en kryss-sammenligning mellom BOPS og CHEOPS for å se at scoringsverdiene for de to skalaene skulle tilsvare samme verdi, ingen smerte (0-2 poeng) moderat smerte (3-4 poeng) og stor smerte (5-6 poeng). $p < 0,05$ ble sett på som statistisk signifikans mellom de to verktøyene. Det var sterk enighet mellom ulike sykepleieres BOPS-scores. Interrater-reliabiliten for ansiktsuttrykk var 0,86 (90%), for verbalisering 0,92 (93%) og kroppsstilling 0,95 (97%). Den totale interrater-reliabiliteten for verktøyet var 0,93 (89%). BOPS og CHEOPS hadde positiv korrelasjon noe

som indikerer at begge verktøyer beskriver de samme variablene. Korrelasjonen mellom BOPS og CHEOPS var statistisk signifikant ($r=0,871$ $p<0,001$). I kryss-sammenligningen ble BOPS delt inn i 3 kategorier. I gruppen uten smerter var median scoren for CHEOPS 6, i moderat smertegruppe var scoren 9 og hos de med store smerter var medianverdien 11. 89% av sykepleierne valgte å gi smertestillende ved BOPS-score >2 . BOPS scoren var høyere (medianscore 5) før smertestillende ble gitt, enn 15 min etter da var verdien 0. Det samme var verdien etter 30- og 60 minutter. Når de sammenlignet BOPS og CHEOPS var likeverdigheten mellom skalaene på 96%. De mente at påliteligheten til skalaen, og at den faktisk målte smerte ble vist ved at observasjonsmålingene fikk en lavere score etter administrasjon av smertestillende. En effekt av smertestillende ble vist i alle barn med BOPS > 2 , og dette styrker anbefalingen om at man skal gi smertestillende dersom pasienten har en scoring >2 poeng. En scoring på 0-2 i BOPS tilsvarer en verdi på 4-7 poeng i CHEOPS.

Score	Facial expression	Verbalization	Body position
0	Natural/positive, smiling, composed	Normal conversation, laugh, crow	Inactive, lying relaxed with all extremities or sitting, walking
1	Negativ facial expression and/or concerned	Completely quiet or sobbing and/or complaining but not because of pain.	Restless movements, shifting position and/or touching wound or wound area.
2	Negative facial expression grimace, disorted	Crying, screaming and/or complaining about pain	Lying rigid and/or drawn up with arms and legs to the body
Utarbeidet tabell fra figur 1 (22) Behavioural Observational Pain Scale (BOPS).			

Alder Hey Triage Pain Score (AHTPS): Er en observasjonsskala som ble utviklet fordi det var misnøye rundt andre tilgjengelige skalaer, og at selvrappotering ikke var tilstrekkelig for alle barn i akuttmottaket. Verktøyet er spesifikt laget for bruk i akuttmottak (29). Skalaen ble utviklet ved hjelp av studier gjort på andre skalaer (CEHOPS og TPPPS), hvor forskerne hadde sett på de ulike variablene hver for seg, og hvor valide verktøyene var for smerte. Skalaen tar for seg 5 ulike variabler: «Cry or voice», «Facial expression», «Posture», «Movement» og «Colour». Hver av variablene har en mulig scoringsverdi 0, 1 eller 2, som resulterer i en totalscore mellom 0-10. Det følger også med en detaljert beskrivelse for de ulike scoringsverdiene. Stewart og hans forskningsgruppe (29) gjennomførte en studie 2002 hvor de undersøkte verktøyet. Skalaen ble testet i akuttmottaket ved at 27 ulike sykepleiere som var trent i å bruke skalaen, observerte 575 barn mellom 0-15 år som kom til

akuttmottaket med ulike smertefulle tilstander. Skalaen ble sammenlignet med WBFPRS for barn >3 år (292 barn). Interrater-reliabiliteten ble undersøkt ved å se om de ulike sykepleierne kom frem til samme scoringsverdi for de ulike barna. For å se om skalaen faktisk målte smerte ble det gjennomført scoringsvurdering før og etter barna hadde fått smertestillende. Tilstandene hos de ulike barna var forskjellige, og de ble delt inn i grupper etter hvor alvorlig tilstand de hadde. De fikk ulike smertestillende avhengig av smertetilstand. Cut-off verdien for interrater reliabiliteten ble satt til «god enighet» dersom den var >0,8. Det var stor enighet i smerteskoringsverdiene mellom de ulike sykepleierne (kappa 0,84). Dette støtter at skalaen kan brukes av trente sykepleiere i akuttmottaket. Det var signifikant forskjell i smerteskoringsverdi før og etter smertestillende medisin $p < 0,001$, som viser at verktøyet kunne skille mellom ulike nivåer av smerte. Assosiasjonen mellom AHTPS og WBFPRS var lavere før ($r = 0,46$, 95% CI 0,37-0,55) enn etter smertestillende medisin ($r = 0,53$, 95% CI 0,45-0,61), og resultatene viste at sykepleierne ga en lavere score enn barna selv. Verdiene kan bære preg av andre faktorer enn smerte, og det er derfor viktig å bruke verktøyet riktig og i en klar kontekst (29).

Response	Score 0	Score 1	Score 2
Cry og voice	No complaint/cry. Normal conversation	Consolable – not talking negative	Inconsolable – Complaining of pain
Facial expression	Normal	Short grimace <50% of time	Long grimace >50% of time
Posture	Normal	Touching/rubbing/sparing	Defensive/tense
Movement	Normal	Reduced or restless	Immobile or thrashing
Colour	Normal	Pale	Very pale/ «green»

The Alder Hey Triage Pain Score; scoring chart. Tabell utarbeidet fra table 1 (29)

Evaluation Enfant Douleur (EVENDOL): Regnes som en ny skala (2012) som har blitt validert for akutt- og mer langvarig smerte, smerter i forbindelse med prosedyrer i akuttmottak og postoperativt for barn i alderen 0-7 år (12). Skalaen ble utviklet av en fransk pediatrik ekspertgruppe, som mente det var behov for å utvikle en smerteskala som kan brukes i alle aldre, som er rask og enkel å bruke i akuttsituasjoner og som kan brukes ved alle typer smerter. Skalaen skulle også gjøre det enklere med oppfølgingsevaluering etter at smertestillende er blitt gitt (56). En gruppe på 291 barn under 7 år ble inkludert i studien. Målekvaliteter som ble undersøkt var 1) indre konsistens evaluert med Cronbach alfa-

koeffisient, 2) interrater-reliabilitet, evaluert av en vektet kappa-koeffisient i to situasjoner 3) konstruksjonsvaliditet bestående av 6 stadier: kontrasterende grupper (a priori med smerte eller a priori uten smerter), følsomhet for endring (før og etter smertestillende behandling), sammenhenger mellom EVENDOL og følgende skalaer CHEOPS, FLACC, TPPS, FPS-R for barn over 4 år og FPS-R og NRS for barn over 6 år .

Korrelasjoner mellom EVENDOL og feber, utmattelse, grad av angst og sult for å vurdere skalaens diskrimineringssevne. Brukervennlighet ble også testet. EVENDOL ble testet ved barnets ankomst og etter smertestillende medisin, i ro og under mobilisering. I utviklingen av skalaen var det første steget å teste ulike variabler på en liten gruppe pasienter for å se om de var representative for små barns smerteopplevelse i akuttmottak. I andre steg ble skalaen testet på en større gruppe pasienter, både med og uten smerter for å undersøke validiteten til verktøyet. Skalaen ble utviklet av en rekke pediatrike spesialister, sykepleiere og metodologer fra fire franske akademiske sykehus. For å velge ut hvilke variabler som skulle være med i skalaen så de på hvilke variabler som var inkludert i tidligere skalaer både akutt (gråt, agitasjon og ansiktsuttrykk), og mer langvarige smerteopplevelse (immobilisering og utmattelse).

Smertelindrende stillinger og beskyttelsesbevegelser blir observert i nesten alle smertefulle sammenhenger, og ble dermed inkludert i den endelige skalaen. Den diagnostiske verdien av kjennetegn som «trøstbarhet» og «interaksjoner med andre og miljøet» viste seg å være helt avgjørende for smertescoren man fikk. Fysiologiske faktorer (hjerterefrekvens og ansiktsfarge) ble ikke inkludert i den endelige skalaen på grunn av deres manglende spesifisitet. Arbeidsgruppen bestemte seg for å bruke 6 kategorier: gråt, trøstbarhet, ansiktsuttrykk, bevegelser, kroppsholdning og interaksjoner med miljøet.

Denne versjonen ble testet på 68 barn, og de beregnet parvis korrelasjon mellom de ulike elementene. «Trøstbarhet» og «interaksjon med miljø» var sterkt korrelert ($r=0,85$, for vurderinger utført i ro, og $r=0,94$ for vurderinger når barna beveget seg). På grunn av den gode korrelasjonen ble disse elementene slått sammen til en variabel som de kalte «interaksjon med miljø». Den endelige versjonen av skalaen inkluderer 5 elementer, hvor hver er beskrevet kortfattet i ord eller uttrykk designet for å unngå tolkningsfeil.

Scoringssystemet består av fire nivåer (0-1-2-3) da forskerne mente at dette var tilstrekkelig for å fange opp all ulik smerteopplevelse i observasjonsperioden. Både intensitet og varighet skulle registreres. Totalscoren var mellom 0-15. De uavhengige observasjonene ble gjennomført av akuttsykepleier og én forsker. Cronbach-koeffisienten var 0,83-0,92.

Konstruksjonsvaliditeten ble demonstrert ved en reduksjon i score etter Nalbuphine 8,14 til 3,62 av 15 i ro ($p < 0,0001$), 11,87 til 6,65 ved mobilisering ($p = 0,0011$). Det var gode korrelasjoner mellom EVENDOL og numerisk score (NRS) til sykepleier eller forsker: 0,79 til 0,92 ($p < 0,0001$). Det var gode korrelasjoner mellom barnas selvrapportering og EVENDOL i alderen 4-7 år (0,64 til 0,93). Diskrimineringsgyldighet i forhold til tretthet, angst og sult var også bra. Interrater-reliabiliteten mellom sykepleier og forsker viste vektet kappa 0,7 til 0,9. Behandlingsterskelen ble bestemt ved 4 poeng av 15 mulige. Forskere mente at de med sine tall kunne stå inne for at EVENDOL har utmerket gyldighet og kan brukes av alle barn under 7 år i akuttmottak, for alle typer smerter; både akutte og langvarige (56).

A pain scale for children under 7

EVENDOL

A pain scale validated for children from birth to 7 years

Score ranges from 0 to 15.

Treatment threshold: 4/15.

Note everything you observe, even if you think the symptoms are not due to pain but to fear, tiredness or illness severity.

Name	sign absent	sign weak or transient	sign moderate or present about half the time	sign strong or present almost all the time	Assessment at admission		Following assessments and/or after analgesic ¹			
					at rest ¹ (R)	during examination ¹ or mobilization (M)	R		M	
							M	R	M	R
Vocal or verbal expression <i>cries and/or screams and/or moans and/or complains of pain</i>	0	1	2	3						
Facial expression <i>furrowed forehead and/or frown, furrowed or bulging brow and/or tense mouth</i>	0	1	2	3						
Movements <i>restlessness, agitation and/or rigidity and/or muscular tenseness</i>	0	1	2	3						
Postures <i>unusual and/or antalgic posture and/or protection of the painful area and/or immobility</i>	0	1	2	3						
Interaction with the environment <i>can be comforted and/or interested in playing and/or interacts with people</i>	normal 0	low 1	very low 2	absent 3						
Remarks	Total /15									
	Date & Time									
	Signature									

Evendol smertescoringsverktøy. Bilde hentet fra kilde (56)

6.3.1 Sammenligning av de ulike observasjonsskalaene

Hesselgard og medarbeidere (22) som evaluerte det nye verktøyet BOPS i forhold til CHEOPS fant ut at CHEOPS er en av de første og mest kjente av observasjonsskalaene, men inkluderer 6 ulike variabler man skal vurdere. Etter at CHEOPS ble utviklet har det kommet

flere og enklere skalaer man kan ta i bruk. Det er i dag økende behov for en enklere og raskere måte å evaluere smerte på, slik at man kan benytte verktøyet på daglig basis, da det ofte er stressende og mye å gjøre på sykehus. BOPS er enkel og rask å bruke noe som gjør den fordelaktig, fremfor de mer avanserte skalaene. Denne skalaen er på mange måter ganske lik FLACC (face, legs, activity, cry and consolability), men enda enklere å bruke da den kun inkluderer 3 variabler, mot FLACC's 5 variabler. Forskerne mente at BOPS viste en overlegen interrater-reliabilitet med en verdi på 0,93. Det var overraskende at det var kroppsposisjonen som hadde den beste påliteligheten. Ut fra andre artikler forskerne hadde sett på, fant de at andre smerteskalaer ikke hadde samme høye interrater-reliabilitet.

For eksempel var verdiene for FLACC 0,52-0,82 og Comfort scale 0,54-0,93. Det viste seg at en forenkling av skalaen ikke gjorde verktøyet mer unøyaktig, hvor man da ville forventet en lavere verdi. Flere faktorer kan påvirke sykepleierens scoringsverdier postoperativt. Rett før barna våker opp fra anestesi kan de være rastløse i en periode før de våkner helt. Sult, frykt og angst samt foreldreseperasjon kan gi feil scoringsverdi i alle observasjonsskalaer (22).

En systematisk oversikt skrevet av von Baeyer og Spagrud (14) anbefaler CHEOPS og FLACC for bestemmelse av smerte forårsaket av medisinske prosedyrer eller andre kortvarige smertegivende hendelser. Begge er mye brukt og har godt dokumentert validitet og reliabilitet (14). En fordel med CHEOPS er at den kun er basert på barnets direkte observerbare oppførsel, uten at det krever vurdering av innsats som trøstbarhet som det gjør i FLACC. FLACC anbefales som førstevalg ved postoperativ smerte på sykehus (14, 27). Grunnen til dette er at den er validert i denne kliniske sammenhengen for en bredere aldersgruppe enn CHEOPS. I tillegg kan variablene som trøstbarhet eller respons på støttende kontakt være mer interessant og gjennomførbart å vurdere i en postoperativ sammenheng (som varer over en lengre tidsperiode) kontra i en smerte forårsaket av en prosedyre som er mer kortvarig.

Observasjonsverktøyene er utviklet for å vurdere intensiteten av «behavioral distress» assosiert med smertefulle hendelser som postoperativt og i akutte situasjoner. Man tar utgangspunkt i barnets atferdsmessige signaler. Ingen spesifikke verktøy er utviklet kun for å vurdere barns kroniske smerter (57). Peden og medhjelpere (83) har sammenlignet en skala kalt The Derbyshire Children's Hospital Paediatric Pain Chart (DPC) som blir brukt ved Derbyshire Children's Hospital for postoperativ smerte, med TPPPS-verktøyet. DPC-skalaen tar for seg parameterne: facial expression, verbal (i.e. selvrapporing) og body movement. Skalaen er ment å kunne brukes av barn i alle aldre i en postoperativ setting. Forskerne mener at skalaen er brukervennlig, ved at den er lett og rask å bruke, slik at den ikke blir en tidsbyrde for

helsepersonellet (83). Barnas smerte scores fra 0-3 (0=ingen smerte, 1= mild smerte, 2=moderat smerte, 3=alvorlig smerte). Hovedparameteren man velger å benytte velges ut fra alder. F.eks. er ansiktsuttrykk den viktigste parameteren hos babyer, mens verbal rapportering vil være mer signifikant hos eldre skolebarn. Selv om alle parameterne scores, er det parameteren med høyest score som legges mest vekt på, og som avgjør hvilke tiltak man skal igangsette. Barna inkludert i studien var mellom 1-5 år og skulle gjennom et kirurgisk inngrep. Alle barna scoret 0 på begge skalaene pre-operativt. Det var en signifikant korrelasjon mellom TPPPS og DPC ($r=0,89$) og forskerne mente at begge verktøyene er like gode. Konstruksjonsvaliditet, konvergent validitet og interrater-reliabilitet ble også målt i denne studien, og hadde høye verdier. Diskrimineringsvaliditet og prediktiv-validitet ble ikke sett på i denne studien, forskerne understrekte derfor at det trengs videre forskning innen forskjellige aldersgrupper, og andre kliniske omgivelser for å støtte verktøyets bruk i klinikken (83).

7. Diskusjon

Systematisk smertevurdering er nødvendig for å kunne gi god smertelindring og vurdere effekten av både medikamentelle og ikke-medikamentelle smertelindrende tiltak. Smerteevaluering hos barn er ekstra utfordrende av flere årsaker. For det første er smerte en subjektiv opplevelse, som ikke lar seg observere direkte. Barn har også ikke fullt utviklede verbale og kognitive evner; dette kan gjøre det vanskelig for dem å uttrykke sitt smertebilde.

Rapporteringen og derfor evalueringen av smerte påvirkes av barnets kognitive funksjon, klinisk kontekst, og smertens typologi. Med økende alder vil barn etterhvert klare å identifisere et smertestimuli, lokalisere det, kvantifisere det, og tilslutt uttrykke smerten verbalt (3). Smerte induserer forandringer i oppførsel, i tillegg til fysiologiske og metabolske forandringer. Det er utviklet et stort antall ulike smerteskalaer for både selvrapportering, og verktøy som baserer seg på observasjonelle oppførsel- og fysiologiske tegn, men kvaliteten på disse varierer (44).

Artiklene funnet i dette litteratursøket tok for seg 8 selvrapporteringsskalaer, hvorav fire var skalaer basert på ansikter (Oucher pain scale, WBFPRS, FPS og FPS-R) (4, 6, 18, 20, 21, 23, 37-49), og fire skalaer basert på numeriske verdier (CAS, RPS, NRS og VAS) (52-62). Det var også artikler som sammenlignet de ulike skalaene med hverandre (5, 7, 10, 15-17, 25, 26, 65, 68, 72). Artiklene tok også for seg 10 ulike observasjonsskalaer (FLACC, CHEOPS, TPPPS, PEPPS, M-PEEPS, OPS, CHIPPS, BOPS, AHTPS og EVENDOL) (2, 3, 12, 14, 18, 22, 27-29, 45, 54, 56, 60, 64, 75, 83, 84, 88), og fem artikler sammenlignet selvrapporteringsverktøy med observasjonsverktøy (14, 82-85).

Helsepersonell må vurdere flere dimensjoner av smerte: 1) smerteintensitet, 2) lokalisasjon, 3) varighet, 4) sensorisk kvalitet, 5) kognitive aspekter ved smerten, 6) affektive aspekter ved smerteopplevelsen, og 7) kontekstuelle faktorer som kan påvirke barnet smerteopplevelse (77). Selvrapporteringsverktøy anses som ønskelig å bruke når dette er mulig utfra pasientens utviklingsnivå og sykdomstilstand (14), men det er også utfordringer knyttet til bruken av disse skalaene. Når selvrapporteringsverktøy er for komplisert for barnet er det helsepersonell som gjør vurderingen ved hjelp av observasjonsskalaer, som baserer seg på atferd og fysiologiske målinger (14). De mest anerkjente og brukte verktøyene er representert i denne litteraturstudien, og nedenfor vil jeg ta for meg ulike utfordringer ved ulike verktøy.

7.1 Selvrappoteringsverktøy

7.1.1 Smerteverktøy basert på ansikter

Tall vs. Bilder: Det finnes studier som har sammenlignet de ulike ansiktsskalaene, og det er listet opp fordeler og ulemper hos dem alle. I en artikkel skrevet av Margo McCaffery (81) har han sett på fordeler og ulemper ved FPS-R og WBFPRS og kommet frem til at en ulempe ved WBFPRS er at tallene som er beskrevet under hvert ansikt ikke nødvendigvis korrelerer med en 0-10 numerisk skala. I studien gjort av Gerra (58), hvor de sammenlignet WBFPRS og VAS, fant de også kun moderat sammenheng mellom skalaene. Disse to studiene kan tyde på at ansiktsskalaer og tallskalaer ikke nødvendigvis korrelerer så godt.

Bieri (50) fant ut at tallene under hvert ansikt i FPS ikke viser god nok sammenheng med numeriske verdier for tallskalaer, og at intervallene mellom ansiktene overlapper. Ansiktene representerer derfor ikke en intervallskala, noe man i utgangspunktet hadde trodd. Carrie Hicks og hennes forskningsgruppe (4) fant derimot en sterk positiv korrelasjon mellom FPS-R og numeriske skalaer 0-10, slik at man basert på denne studien mener at skalaen kan oversettes til numeriske verdier (4, 81).

Ønsket om å kunne oversette ansiktsskalaene til bestemte tallverdier er basert på argumentasjon om at dette gjør det enklere for helsepersonell å tolke resultatene og ut ifra dette vite hvordan de skal behandle smerten (23). De sprikede resultatene gjenspeiler nødvendigheten av mer forskning på området. De nye computerbaserte skalaene som er under utvikling tar for seg dette problemet, og et av målene med disse er å gjøre det enklere for helsepersonell å tolke resultatene (23).

Kontinuerlig variabel: Et problem med de tradisjonelle ansiktsskalaene, som også delvis er beskrevet i avsnittet over, er at de ikke nødvendigvis representerer en kontinuerlig variabel som forandres systematisk ettersom smertefølelsen avbildes fra ytterkantene «no hurt» til «worst hurt». Det diskuteres derfor om smerteintensitet er målbar på denne type skalaer, og om man i det hele tatt kan tilskrive hvert enkelt ansikt noen objektiv verdi (23). WBFPRS er også asymmetrisk ved at det er to smilende ansikter, ett nøytralt ansikt og tre triste ansikter. Forskere som har utviklet skalaen mener å ha antatt at den subjektive forskjellen mellom hvert par av tilstøtende ansikter er den samme, og at det hele derfor blir en intervallskala som tillater bruk av parametriske statistikk (23). Dette er likevel ikke forskningsmessig fullt understøttet (23, 24).

Affektiv mimikk vs. smerte: Det har vært diskutert om inkludering av smil og tårer på ansiktene kan føre til a) forvirring ved at barnet tror at han/hun må gråte når de har vondt b) signifikant høyere smerteverdier gitt på skalaer som har smilende ansikter for å representere «ingen smerte» sammenlignet med skalaer som har nøytrale ansikter som uttrykk «ingen smerte» (4). En mye diskutert potensiell feilkilde med WBEFPS er at ansiktet som representerer «no hurt» og «hurts little bit» er smilende, som kan gjøre at barnet feiltolker og dermed tror at han/hun må velge «hurts a little more» ansiktet hvis de skulle ha mild smerte fordi de selv nødvendigvis ikke smiler. Samme problem gjelder ansiktet «hurts worst» som gråter. Dette kan få barnet til å tro at han/hun ikke kan velge dette ansiktet med mindre han/hun gråter. Ved bruk av denne skalaen er det derfor viktig at man informerer barnet om at de skal peke på ansiktet som representerer hvor vondt de har, og ikke hvordan ens eget ansikt ser ut (81).

Bieri og hans forskningsgruppe (50) forsøkte å endre på dette ved utviklingen av FPS. De mente at en fordel med denne skalaen er at ansiktene ikke skildrer smil eller tårer, noe som reduserer potensiell forvirring knyttet til ansiktene (17, 50). Et annet diskutert potensielt problem med ansiktsskalaer er at antall og type egenskaper som er avbildet i hvert enkelt ansikt ikke alltid er like. Det er derfor ikke én enkelt forandring i ansiktet som er avbildet som kan knyttes til endringen i smerteintensitet (23).

Språk og etnisitet: FPS-R skalaen er oversatt til mer enn 32 språk, noe som gjør den brukervennlig for barn og helsepersonell fra ulike land (17). For at slike oversettelser skal anses som gyldig har de blitt testet på grupper av barn fra de ulike landene. Den teoretiske koblingen mellom kultur og smerte er basert på ideen om at barns uttrykk for smerte og sykdom læres av foreldre, familie og andre mennesker innen samme kultur. Kultur kan forklares som et rammeverk for læring av atferd og kommunikasjon som er med på å forme barnets perspektiv på helse og sykdom. På denne måten vil foreldrenes tro, standarder og forventede normer forme barnets forventning av smerte og valg av mestringsstrategier (59).

Det er usikkert hvilke faktorer helsepersonell bør være oppmerksomme på når de skal evaluere smerte hos barn fra ulike kulturer. Et første skritt for å ta tak i mange av disse spørsmålene er å tilby mer valide og kulturfølsomme smerteverktøy. Bruken av disse verktøyene vil hjelpe helsepersonell med å integrere kultur i praksis og kan åpne nye dører for forskningsspørsmål om kulturens innflytelse på smerter. En utfordring ved tolkning av scoringsverdier i ulike kulturer er å skille mellom barnets respons påvirket av selve kulturen

og barnets evner til å forstå en selvrangeringsskala (59). Skalaene må derfor gå igjennom nøyaktig testing for ulike etniske grupper før de kan benyttes.

Oucher pain scale er et av verktøyene som er ment å være kulturfølsom, og de ulike versjonene sies å ha god validitet og reliabilitet (48). Det er forventet at skalaer som tar hensyn til kultur vil kunne bedre kommunikasjonen mellom helsepersonell, barn og pårørende med ulik kulturell bakgrunn (74). Siden ansiktsskalaen er basert på bilder eller tegninger er språk ikke en nødvendighet for å kunne bruke skalaen. Disse skalaene kan være nyttige hos barn som ikke er godt språkutviklet, har språkproblemer, eller har annet morsmål enn helsepersonellet.

Kjønn: Flere av studiene fant at kjønn har lite å si for utfallet av scoren (58, 61). Abu-Saad og kolleger (43) fant derimot at jenter velger mer komplekse ordbeskrivelser enn gutter når de skal beskrive smerte. Dette kan reflektere oppdragelsen av barna. Generelt i samfunnet har man også en forestilling om at jenter er mer åpne enn gutter. Kanskje påvirker denne forventingen barn allerede tidlig i livet. Det viste seg også at jenter generelt ga høyere smertescore enn gutter. I en annen studie gjort av Gharaibeh og Abu-Saad (59) fant de at jenter og gutter foretrakk ulike smerteverktøy. I valget mellom Poker sjetong verktøyet, ansiktsverktøy og ordbeskrivende verktøy foretrakk jentene Poker sjetong verktøyet, mens guttene likte best ansiktsskalaen (59). Det var ingen forskjell mellom aldersgruppene. Denne studien undersøkte barn fra Jordan, og funnene var altså annerledes enn studier gjort på barn fra vesten (5, 74), hvor kjønn hadde lite å si for hvilket verktøy barnet foretrakk å benytte.

Disse funnene gjenspeiler viktigheten av kultursensitive skalaer, som beskrevet i avsnittet over. Arabiske barn sosialiseres til å leve innenfor forventningene til de kulturelle normene, verdiene og kjønns-spesifikke rollene. Kanskje er dette med på å påvirke hvordan barna uttrykker smerte (59). Akkurat som skalaene er kultursensitive bør de også være kjønns-sensitive. Oucher scale har fått kritikk for kun å avbilde gutter. Utviklingen av kjønnsfølsomme skalaer ser ut til å styrke jenters evner til å identifisere seg med fotografiene; dette kan styrke validiteten til skalaen (74).

Alder og kognitive bias: Det er stor diskusjon rundt i hvilke aldersgrupper selvrangeringsverktøy kan benyttes, og det er lite forskning på bruk av selvrangeringsverktøy for barn yngre enn 5 år (79). Mange barn under 6 år har i studier vist at de ofte prefererer til å score topp og/eller bunnverdiene på skalaene, såkalt kognitiv bias (16, 79).

Det kognitive biaset å velge enten topp- eller bunnverdier skjer spesielt hos 3-åring, 4-åring og noen 5-åring (6). Mønsteret beskrives som et «anchor bias», som er et eksempel på en «respons bias». Begrepet beskriver en strategi barn bruker for å kunne svare på et spørsmål selv om de egentlig er ute av stand til å forstå hva spørsmålet betyr. Eksempler på slik bias kan være at respondenten velger å konsekvent si «ja» som svar, velger runde tall, eller velger sitt favoritt-tall. En annen type respons bias er «sekvens bias» som kan oppstå ved bruk av ansiktsskalaer: Barnet velger for eksempel ansiktet helt til venstre på det første spørsmålet, og velger deretter ansiktet til høyre for dette ansiktet ved hvert av de andre spørsmålene som blir stilt. Man får da en sekvens av scoringer, enten oppover eller nedover (0-2-4-6-8) (16, 79).

Carl L von Baeyer og medarbeidere (79) studerte responsbias hos 185 barn, 3-5 år gamle, ved bruk av to bildeskalaer (CPPP og FPS-R). Charleston Pediatric Pain Pictures (CPPP) er et verktøy med 17 ulike bilder av hendelser som skal representere alt fra ingen smerte til mye smerte. Barna skulle anvende FPS-R for å angi hvor mye smerte bildene representerte. De fant at forekomsten av responsbias falt med økende alder. Studier som bruker FPS-R indikerer at 5-åring og de fleste 4-åring kan bruke dette verktøyet på en gyldig måte, mens skalaen er ganske kompleks for de fleste 3-åring og valideringsdata mangler for denne aldersgruppen.

Et enklere smerteverktøy for 3-åring (17, 59, 79) er «Pokerchip tool» som benytter 4 poker sjetonger, som hver representerer samme mengde smerte (jo flere sjetonger, jo mer smerte). Barnet blir spurt hvor mange sjetonger smerten deres representerer. Antallet sjetonger blir brukt som mål i stedet for ansiktsskalaen, noe som gir et mindre antall respons-alternativer (4 istedenfor 6). Dette har vist seg å være fordelaktig for mindre barn (17). Poker sjetong-verktøyet har vist korrelasjon med andre selvrapporteringsverktøy som FPS-R, Oucher-skalaen, VAS og WBFPRS, og også observasjonsskalaer (71). Wong and Baker sammenlignet i 1998 pålitelighet og validitet av 6 smerteverktøy for barn, og fant at Poker sjetong-verktøyet hadde størst validitet (59). Poker sjetong-verktøyet har blitt anbefalt som bruk på 3-åring (59), men gode nok studier for barn på 3 år er også sparsomme for dette verktøyet. Quinn og medarbeidere (70) tenkte at siden det viste seg at de minste barna ofte prefererte til ytterpunktene av skalaene, ville det muligens være en løsning å redusere antall ansikter fra seks til tre, slik at skalaen ble mindre kompleks. Tanken støtter Baeyer og kollegaer, som fant at det kan være fordelaktig med færre responsalternativer for de minste barna (79). Evidensen for

valg av selvrapporteringsverktøy for smerte hos barnepasienter er på nåværende tidspunkt generelt ikke godt nok (16, 17).

Klinikere bør være klar over at de minste barnas smertescore kan være misvisende, spesielt når smerteskalaen bare skal brukes én gang for å måle smerter ved én enkelt anledning. Dette gjør det ekstra vanskelig å oppdage underliggende responsforskjeller. Selvrapporteringsresultater fra barn under 5 år bør generelt behandles med forsiktighet og bekreftes med observasjonelle skalaer. Avvik kan deretter utforskes nærmere ved observasjon eller diskusjon med foreldrene (79).

Abu-Saad og kolleger (91) fant i sin studie at modning og tidligere erfaringer av smerte, gjør at barna bruker mer komplekse ord i sin smertebeskrivelse. Ordene barna velger for å beskrive smerte gjenspeiler i stor grad deres tidligere opplevelser og erfaringer. Yngre barn ga generelt høyere score enn eldre barn, noe som kan skyldes at de ikke har bred erfaring med smerteopplevelser enda. Det viste seg at barn som hadde mye erfaring fra sykehus rapporterte lavere smertescore (43).

Papirbasert vs. Computerbasert: Franciullo og medarbeidere hevder at en ulempe ved papirbaserte ansiktsskalaer er at de ikke tillater at barnet kan forandre og finjustere verdien av smerte før de bestemmer seg for en endelig verdi (23). Et annet problem er at statiske ansiktsmodeller ikke tillater registrering av flere endringer i scoring over tid før barnet bestemmer seg for hvilket ansikt de skal velge. Det finnes ingen funksjon i papirbaserte metoder som gjør at et barn kan modifisere eller rette opp en første bestemmelse, utover bare å fortelle administratoren at de ønsker å ombestemme seg.

Et ytterligere problem er at registreringer knyttet til skalaene ikke automatisk lagres i en datafil for senere analyse (23, 24). En ny computerbaserte ansiktsskala er utviklet som en målemetode hvor man får en symmetrisk og kontinuerlig skala. Ved denne metoden justerer barnet formen på munnen, samtidig som de kan skifte øynene fra åpne til lukkede. Endringene i munn og øyene er korrelert, slik at høyden på munnen er skalert fra -10 (sur) til +10 (smil). Verdien og det digitale bildet av ansiktet lagres automatisk for senere evaluering av klinikerens. Fordelene med et slik system er at barnet kan velge ansikter fra et bredt spekter, som gjør det mulig å vurdere en rekke smertetilstander. Klinikerens sparer også tid ved at de slipper dataregistrering, da bildene lagres i databasen automatisk. I tillegg kan det utvikles et system som gjør at resultatene kan leses av, så lenge man har tilgang til pasientens journal via

en PC (23, 24). Studiene rapporterte at flertallet av barna at foretrakk å bruke computer-versjonen fremfor papirversjonen av ansiktsskalaer (23, 24, 26). Scoringer med computerversjon var signifikant korrelert med Wong and Baker Faces scale. Dette betyr imidlertid ikke nødvendigvis at det var et godt samsvar i scoringsverdiene, kun at begge skalaene representerer stigende smerte (24).

I en annen studie gjort av Terri Sun og medarbeidere (26) hvor de sammenlignet computer-basert skala av FPS-R med papirversjonen fant de sterk sammenheng mellom papirutgave og computerversjon, FPS-R ($r > 0,96$) for barn i alderen 4-7 år. Basert på denne studien er computerbasert og papirutgave av ansiktsskala tilnærmet like gode.

En ulempe med de computerbaserte skalaene hvor barn tegner ansiktet de selv mener representerer deres smerte, er at tegningen kan påvirkes av hvilket humør barnet er i. Det stilles derfor spørsmål ved om barna klarer å bruke verktøyet kun for å beskrive smerte, eller om andre faktorer påvirker tegningen. Dette kan f.eks. være at de føler seg ulykkelige over at de må være på sykehus, og at de kjeder seg. Disse følelsene kan være vanskelig å skille for et barn, da de ofte henger sammen. Grunnen til at det er viktig å skille mellom smerte og ulykkelighet er at et barn med smerter sannsynligvis må behandles med smertestillende, mens et barn som er ulykkelig ikke nødvendigvis trenger smertestillende men kanskje sosialisering og noen å prate med. De fleste barn på sykehus vil nok i mange tilfeller ha behov for begge deler (24).

7.1.2 Smerteverktøy basert numeriske verdier og farger

Alder og kognitiv bias: Seriering, som er prosessen hvor man organiserer objekter i en ordnet serie, for eksempel i økende størrelse eller høyde, er en viktig kognitiv forutsetning for å kunne bruke tallskalaer som f.eks. VAS riktig. VAS krever at barn skiller oppfatningene sine fra små til store. Denne evnen vises vanligvis ikke før barnet er 7 år (86). Selv om det er stor aldersvariasjon for når barn har utviklet serieringsevne, antyder dette generelt at de fleste 5-åringer ikke vil kunne bruke en VAS-skala konsekvent. Det er likevel studier som har rapportert gyldig og pålitelig bruk av VAS helt ned i 5 års alder (4, 61). Siden det er stor variasjon i forhold til alder og evne til seriering er kanskje den kognitive evnen en bedre prediktor for et barns vellykkede bruk av VAS, enn kronologisk alder.

Kognitiv evne blir ofte vurdert ved å administrere en standardisert intelligensstest, som gir en poengsum, intelligenskvotienten (IQ). Imidlertid påvirker flere faktorer en persons ytelse på standardiserte IQ-tester, inkludert sosioøkonomisk status, kjønn, og verbale-, romlige og

matematiske ferdigheter (86). Om testen gjøres i rolige eller stressende omgivelser har også noe å si. Man kan for eksempel forvente at barna vil ha større vanskeligheter med å gjennomføre VAS-skalaer riktig i stressende situasjoner som f.eks. i et akuttmottak, enn i roligere og kjente omgivelser, som for eksempel på skolen (86).

VAS er oftest horisontal, men har også blitt presentert som en vertikal skala, eller sjeldnere som en kromatisk skala som f.eks. Color analog scale (CAS). VAS-skalaen er en av skalaene som er mest forsket på. Den viser labilitet over en 2 ukers periode, men den viser også god korrelasjon med FSP-R og Oucher. VAS-scoren er også sensitiv for forandringer i smerteintensitet, som for eksempel etter et operativt inngrep (17). NRS har fordelen av at den kan benyttes enten visuelt eller verbalt (VNRS). Det er stilt spørsmål ved om versjonene kan benyttes for ulike aldersgrupper, dvs. om én av versjonene er enklere å forstå enn den andre. Dette spørsmålet krever videre forskning (69, 76).

Hvis barnet skal bruke numeriske selvrapporteringsverktøy må barnet ha evne til numerisk resonering, klassifisering, og enkle mattekunnskaper i form av enkel addisjon og subtraksjon. Det er ifølge Champion (67) ikke nok å bare kunne telle. Argumentene har ført til stor debatt rundt hvilken alder barn er i stand til å bruke numeriske selvrapporteringsverktøy.

Rainbow Pain Scale, RPS, er ment for de aller minste, og skal ifølge oppfinneren være enklere enn både ansikt- og numeriske skalaer. RPS har imidlertid fått kritikk for at dersom barn tidligere har en erfaring med fargene i regnbuen kan dette påvirke fargevalget, slik at barnet ikke nødvendigvis velger en farge basert på sitt uttrykk for egen smerte. Det er en forutsetning før man benytter skalaen at det informeres om at barna skal plukke ut sin egen farge, og at fargen ikke trenger å samsvare med fargene i den ordentlige regnbuen. En annen forutsetning er at barnet forstår denne informasjonen (66).

Språk og etnisitet: De numeriske skalaene har mindre problemer knyttet til språk enn ansiktsskaler med ord beskrivelse. Årsaken til dette er at tall i mindre grad enn ord påvirkes av språkforskjeller. Det er likevel en forutsetning at barnet har evner til å forstå hvordan skalaen skal brukes. VAS som er den mest kjente tallskalaen er med suksess brukt i både Italia og Frankrike (21).

Kontinuerlig gradering: En fordel med de numeriske skalaene som VAS, NRS og CAS er at de er sensitive for endringer i smerte. Det at man kan gradere smerten ut fra en skala, gjør det enklere for pasienten å uttrykke smerteintensiteten mer nøyaktig enn ved en skala hvor

det er et «enten/eller» svar. NRS er mindre sensitiv for endringer enn VAS, noe som muligens kan skyldes at den kun graderes fra 0-10, mens VAS ofte fra 0-100 (21). VNRS har også vist lavere sensitivitet enn blant annet VAS i korte tidsperioder. En mulig forklaring kan være at barna husker tallet de oppga ved forrige måling. En annen ulempe med VNRS er at svarene oppgis i hele tall, kontra en visuell gradering med kontinuerlig finglydert tallinje.

En fordel med VNRS kontra VAS er at VNRS ikke krever noe ekstra utstyr, og kan benyttes raskt og enkelt verbalt. Det er også mulig å gjennomføre en slik skala for eksempel under en telefon-konsultasjon. Siden smertebehandling kan være effektiv uten å nødvendigvis fjerne all smerte, er det ikke tilstrekkelig med en todelt skala som sier kun «smerte» eller «ikke smerte». En gradert visuell skala vil på en mer nøyaktig måte kunne kvantifisere smerten og også gi et mer nøyaktig mål på effekt av den smertestillende behandlingen (89).

Forenkling av smerte: Å gjengi smerteopplevelse kun i form av et tall som skal reflektere intensiteten er en grov overforenkling. Smerte har mange andre egenskaper i tillegg til intensitet, som blant annet lokalisasjon, sensoriske kvaliteter, kognitiv vurdering og affektive reaksjoner. Det er viktig å sette scoringsverdien inn i en klinisk kontekst, hvor man også stiller oppklarende spørsmål ved hjelp av en smerteanamnese. Anamnesen kan inkludere spørsmål som: «når startet smerten?», «er det noe som gjør smerten bedre eller verre?» evt. skademekanisme. Ved kroniske smertetilstander er det viktig å utforske hvordan smerten påvirker daglig funksjon og søvn. Det er viktig med en bred vurdering for at man som helsepersonell skal kunne informere, evaluere og forbedre smertebehandlingen (16).

7.2 Observasjonelle eller objektive smerteverktøy

Alle observasjonsverktøyene bygger på de samme prinsippene, nemlig at en observatør tolker barnets smertegrad ut fra dets oppførsel, og i noen tilfeller fysiologiske målinger (54, 78).

Litteratursøket ga treff på mange observasjonsverktøy. De fleste av disse benytter de samme parameterne, som i hovedsak er «facial expression», «cry» og «body position» i en eller annen form. Enkelte skalaer har ytterligere parametere som kan være tilstede når et barn har smerter.

FLACC, CHEOPS, TPPPS, OPS, PEEPS og M-PEEPS er alle utviklet for at helsepersonell skal kunne evaluere postoperativ smerte hos barn, men i noe ulike aldersgrupper. Alle verktøyene tar for seg de tre parameterne beskrevet over. CHIPPS og BOPS spesifiserer i mindre grad hva slags smerte de kan benyttes for. AHTPS er spesifikk utviklet for bruk i

akuttmottak. I og med at alle skalaene tar for seg de tre parameterne «facial», «cry» og «body position» kan dette tolkes som at disse parameterne erfaringsmessig gir et godt uttrykk for smerte.

Alder: Observasjonsskalaene regnes i hovedsak å kunne brukes fra noen måneders alder frem opp mot ca. 10 års alder. Forskerne har angitt anbefalt alder ut fra hvilke aldersgrupper verktøyene har blitt testet på. Etersom variablene i skalaene stort sett er de samme kan det være et interessant spørsmål, og en mulig hypotese at alle disse skalaene egentlig kan benyttes for de samme aldersgruppene.

Det skjer en endring i smerteatferden med alderen, som f.eks. at eldre barn gråter mindre enn yngre barn tross den samme fysiske smertestimulansen. Mulig kan tidligere erfaringer og kognitive ferdigheter være med på å påvirke dette. Kanskje øker også smerteterskelen med alderen (55). Endringen i hvordan smerteuttrykket forandres med alder kan man tenke seg vil påvirke utfallet av scoringsverdiene. Skalaene tar per i dag ikke hensyn til dette.

De fleste observasjonsskalaene ble opprinnelig designet for å imøtekomme et bestemt aldersområde, men har senere blitt brukt for et bredere aldersspenn. F.eks. ble CHEOPS opprinnelig utviklet for barn i alderen 1 – 7 år, men har vellykket blitt brukt i studier fra 1 mnd - 17 år (14). Det trengs videre forskning for å avgjøre om det er noen gevinst i å justere observasjonsskalaene med tanke på endring i atferd med økende alder, eller om det kan være nyttig å utvikle separate verktøy for forskjellige aldersgrupper (14).

EVENDOL som er nytt verktøy utviklet i 2012 forsøker å bøte på problemet at det er (for?) mange skaler å velge mellom. Forskerne har forsøkt å sammenholde alle de ulike parameterne og inkludere dem i samme verktøy. De mener dette vil gjøre det enklere for helsepersonell, fordi at alle skalaene vi har i dag kan skape forvirring rundt når man skal benytte de forskjellige (56).

Mange skalaer, uten klart definerte grenser for når de skal velges, kan være forvirrende, og muligvis en årsak til at helsepersonell velger å ikke benytte verktøy i det hele tatt.

EVENDOL-skalaen har blitt angitt å være egnet for både akutte, kroniske, og postoperative smerter i aldersspennet 0-7 år. Skalaen er anvendt i Frankrike, Sveits og Belgia, og spanske og nederlandske oversettelser er produsert. Den umiddelbare validiteten (Face validity) av skalaen er likevel enda ikke systematisk testet, og flere studier er nødvendig for å sikre at verktøyet kan tas i bruk til alle formålene det er utviklet for (56).

Objektiv vurdering: En stor utfordring ved observasjonsskalaer er kravet om at travle akutt-sykepleiere skal evaluere smerter raskt, men samtidig riktig, hos et engstelig eller panikkfylt barn hvor mange andre faktorer enn smertebehandling krever oppmerksomhet. Smertevurderingen er ofte avhengig av en kort adferdsanalyse fra sykepleiere. Sykepleierens erfaring, kunnskap, empati, kjennskap og relasjon til barnet er viktige faktorer som vil påvirke evalueringen. Barn har ulike uttrykksformer for smerte. Gråt er et vanlig uttrykk som ofte forbindes med at barn har det vondt. Utfordringer ved observasjonsskalaer er at variabler som er knyttet opp mot gråt kan tolkes som smerte og indusere falsk positive målinger. Likevel er det falsk negative målinger som er mest bekymringsverdige, og trolig det vanligste problemet (64). Noen barn trekker seg tilbake og blir tause, mens andre kan sovne av utmattelse. Smerte hos barn som ikke klager og/eller er uten bevegelse er vanskelig å evaluere med observasjonsskalaer. Scoringer basert på observasjoner av barnets oppførsel kan på denne måten være villedende. Et annet potensielt problem kan være at helsepersonell (mulig ubevisst) kan stole mer på hva de selv observerer enn informasjonen barnet gir om de subjektive opplevelsene (78). I studien gjort av blant annet Stewart (29) hvor de sammenlignet WBFPRS og AHTPS i akuttmottak fant de at sykepleierne oppga en lavere verdi enn barna selv før smertestillende ble gitt. Dette kan indikere at WBFPRS er dårlig egnet i akutte situasjoner, eller at sykepleierne undervurderer barnas smerte. Ansiktsuttrykk er et av variablene som har blitt omtalt som mest fremtredende for evaluering av smerte hos barn ved hjelp av observasjonsskalaer (18, 27). I en studie hvor CHEOPS, CHIPPS, FLACC og OPS ble evaluert viste det seg at skalaene hadde høy spesifisitet men kun moderat sensitivitet, og at smerten hos enkelte barn ble underestimert trolig grunnet at barna hadde en atypisk oppførsel i forhold til hva som forventes ved smerter. Ca. 1 av 10 barn fikk underestimert sin smerte, og den største risikofaktoren så ut til å være dersom barna var stille (27).

Diskriminering av andre faktorer: De fleste smerteskalaer basert på observasjon inneholder elementer som kan tolkes som indikatorer for smerte. Problemet er at de også kan være indikatorer for andre negative følelser, som blant annet frykt og angst. Gråt, sutring, fysisk anspenthet, klamring, rastløshet, ønske om å unngå berøring kan alle observeres både ved smertefulle og ikke-smertefulle tilstander. Få forskere har presentert diskriminerende gyldighetsdata som viser at deres observasjonsskala kan skille smerteintensitet fra andre negative emosjonelle tilstander (14).

Type smerte: Det er diskusjon mellom forskere om når de ulike verktøyene egner seg best for å evaluere smerte. Et sterkt smertefull stimulus gir en øyeblikkelig rask, intellektuell

atferdsrespons i form av tilbaketrekking (f.eks. hvis man brenner seg), vokalisering og grimaser i ansiktet. Denne åpenbare responsen dempes raskt (i løpet av minutter eller timer) i møte med kontinuerlige eller kroniske smerter, og kan erstattes med mer skjulte atferdsmønstre som stivhet, at barnet blir mer stille, og avvergestillinger (14). Ved langvarige akutte og kroniske smerter er fravær av synlige tegn på smerte vanlig, bortsett fra under forverring som oppstår når pasienten f.eks beveger seg (14). Endringen i smerterespons over tid er også noe Beyer og McGrath (78) fant; de så at CHEOPS-score generelt var lavere noen timer etter kirurgi, og da ikke korrelerte godt med egenrapportert av smerte. Dette kan tyde på at variablene i disse skalaene muligens ikke er representative for mer langvarig smerte (87).

Tidligere har det vært dokumentert at FLACC kun kan benyttes ved prosedyre- og postoperativ smerte, men i 2017 ble det gjennomført en studie hvor forskerne også konkluderte med at verktøyet hadde høy reliabilitet og sensitivitet for akutt smerte som ikke var relatert til en postoperativ setting eller prosedyre (64). Strout hevder at M-PEEPS er psykometrisk forsvarlig å bruke til måling av smerte hos pasienter i akuttmottaket, og at skalaen egner seg godt for å evaluere effekten av smertestillende tiltak, og å følge endringer i pasientens smertetilstand over tid (3). Det er generelt dårlig dokumentert om verktøyene kan benyttes i andre settinger enn postoperativ smerte. Observasjonsverktøyene for kortvarig akutt smerte (f.eks. prosedyresmerter) er mye bedre etablert enn metoder for måling av langvarig eller kroniske smerter. Kroniske smerter vil sannsynligvis manifestere seg som mer komplekse forandringer som økt irritabilitet, dårligere humør, søvnvansker, endring i matlyst og skolepresentasjoner. Alle disse fenomenene krever bredere kunnskap om barnets grunntilstand og temperament, og kan være vanskelig for en sykepleier som ikke kjenner barnet å vurdere på egenhånd (3, 86).

7.3 Generelt om smerteverktøy

Sammenligning over tid kontra mellom individer: Generelt er smertescoringsverktøy mer nyttig for sammenligninger over tid og innen samme individ, kontra mellom individer (16). Det kan f.eks. være to barn som har gjennomgått nøyaktig samme prosedyre, men det ene barnet angir smertescore 10/10 og det andre barnet 7/10. Det kan ikke antas at det første barnet har mer smerter enn det andre. Ulikhetene kan skyldes forskjeller i forståelse av smerteskalaen eller ulik tolkning av hva øverste nivået i smerteskalaen representerer. Dersom

barna viser en reduksjon i smertescore over tid og f.eks. etter administrering av smertestillende er dette det viktigste, likesom at en økning i smertescore resulterer i en tettere oppfølging og aggressivitet av behandling (16).

Et annet problem knyttet opp mot individuelle forskjeller er hvor man skal sette cut-off verdi for når pasienten har behandlingseffekt. Hvis et barn får redusert sin smerte fra 8/10 til 6/10 etter intervensjon, kan man da si at dette er en betydelig forbedring og behandlingssuksess? Størrelsen på en klinisk signifikant forandring kan være ulikt mellom barn. Det er ingen standardiserte metoder som har blitt utviklet for å skaffe denne informasjonen eller for å integrere den med den oppnådde smertescoren. Et alternativ som har blitt brukt til minimum signifikant forskjell i kliniske studier er å definere behandlingssuksess som en 30%-50% reduksjon i smerte. Med en 50% reduksjon vil da en reduksjon fra 10/10 til 5/10 og en reduksjon fra 2/10 til 1/10 være likeverdige og betraktet som suksess. I klinisk bruk vil kanskje en vel så god metode være å spørre barnet hvilket nivå av smerte som vil være akseptabelt eller mye bedre, og dette kan da være et realistisk behandlingsmål (16).

Hvilken skala foretrekker barna: I studien som ble gjennomført av Castarlenas og hans forskningsgruppe (52) ønsket de også å finne ut av hvilken skala av NRS, FPS-R, CAS og VAS barna foretrakk å bruke. De fant at den mest foretrekkende skalaen var CAS, mens den som var minst foretrukket var VAS. NRS, var nest mest foretrukket i alle tilfeller, og de var ingen forskjell mellom jenter og gutter. Det var sterk korrelasjon mellom NRS og FPS-R. Dette mente forskerne var positivt ettersom FPS-R er en anerkjent skala som ses på som den nåværende optimal skala for måling av smerte hos pediatriske pasienter (52). Andre studier som også har sammenlignet FPS-R og NRS, har vist at barn oftest foretrekker å benytte FPS-R fremfor NRS. Dette gjelder spesielt barn i yngre aldersgrupper og jenter (67). Årsakene til at barna foretrakk FPS-R fremfor NRS var blant annet at FPS-R var lettere og morsommere og bruke.

Hva bør gjøres videre? Forbedrede selvrapporteringsverktøy og klarere avgrensning av i hvilke aldersgrupper disse kan benyttes krever mer forskning på området. Muligens kunne forenklinger av ansiktsskalaene og bedre bruker-instruksjoner heve kvaliteten. Klinikere bør være oppmerksomme på at småbarns smertescore kan være misvisende, spesielt når smerteskalaen kun brukes én gang for å måle smerte ved én enkelt anledning. Dette gjør det vanskelig for klinikeren på oppdage underliggende responsforstyrrelser. Derfor bør selvrapporterende smerteverdier for barn under 5 år generelt behandles med forsiktighet, og bekreftes av observasjonsmålinger. Det kan for eksempel være nødvendig at sykepleieren

gjennomfører en FLACC og ser om dette samsvarer med barnets selvrapporing på FPS-R. Avvik kan deretter undersøkes videre ved observasjon eller diskusjon med foreldre (79).

En artikkel skrevet fra Alexandra Beltramini (12) tar for seg ulike smerteverktøy, både selvrapporeringsverktøy og observasjonsverktøy, har kommet frem til en egnet cut-off alder ved 6 år. Barn over 6 år kan benytte selvrapporeringsverktøy (også numeriske), mens barn under 6 år skal evalueres med observasjonsverktøy. Siden det er individuelle forskjeller i kognitiv utvikling ut fra kronologisk alder må man regne med at slike cut-off verdier vil kunne gi både falsk positive (ugyldige poeng fra barn som ikke forstår skalaen) og falsk negative (ikke å få gyldige poeng fra barn som forstår skalaen, men som ikke ble spurt) resultater. «The pediatric initiative on Methods, Measurement, and Pain Assessment in Clinical Trials» dokumenterte etter en omfattende litteraturgjennomgang at selvrapporing ikke er tilstrekkelig for barn under 4 år, og at man derfor må gjennomføre observasjonsmålinger av disse (64). Det finnes ingen gode validerte metoder for å identifisere om de minste barna kan eller ikke kan benytte selvrapporeringsverktøy. Det trengs derfor videre forskning og undersøkelser på området.

Ut fra forskningen vi har i dag kan det virke som det er utviklet for mange skalaer, slik at det blir vanskelig å velge hvilken man skal benytte. Alle skalaene har styrker og svakheter. En praksisendring som bør gjennomføres ifølge Quinn og hans forskningsgruppe (70) er å standardisere bruk av mer enn ett smerteverktøy. Mange sykehus og andre helseinstitusjoner velger seg ut én bestemt smerteskala for hver aldersgruppe. Studien viser at flere typer verktøy bør være tilgjengelig for bruk. Institusjonen bør gi veiledning og opplæring om valg av riktig verktøy for akkurat dette barnet.

Ytterligere forskning er nødvendig for å finne ut om sykepleiere kan estimere rapporteringsbias hos enkeltbarn for de ulike skalaene. Sykepleierne trenger mye trening for å kunne bestemme hvilket verktøy som er best egnet for hver enkelt pasient. Mer nøyaktig smertevaluering vil sannsynligvis resultere i mer målrettet smertehåndtering og behandling av høyere kvalitet. Kanskje er løsningen én app med alle de ulike skalaene som har gjennomgått pålitelig forskning (70). Siden smerte er så komplekst virker det sannsynlig at smerten ikke kun kan vurderes ut fra én skala, men at man må ha ulike aspekter for en mer helhetlig vurdering. Foreldrene er de som kjenner barnet best, og de bør derfor også involveres i smertevurderingen. Smertevurdering bygger på god kommunikasjon mellom barn, foreldre og helsepersonell og bør baseres på en bred kartlegging, gjerne i tverrfaglige team.

8. Oppsummering og konklusjon

Smertepåvirkede pasienter er noe en lege møter i de fleste av sine konsultasjoner, og smerter er en vanlig årsak til at folk oppsøker lege. Evaluering av smerte for den pediatriiske pasientgruppen er vanskelig fordi de ikke har fullt utviklede kognitive og verbale ferdigheter og derfor uttrykker smerten på en annen måte enn voksne. Smertevurdering i de yngste aldersgruppene avhenger av observasjonsevnen og ekspertisen til helsepersonell.

Det er utviklet et stort antall smertescoringsverktøy for barn, og disse inkluderer både selvrapporteringsverktøy (ansiktsskaler, numeriske skalaer og fargeskalaer), og observasjonsverktøy. Sistnevnte baseres på tolkning av smerterelatert atferd og fysiologiske målinger gjort av helsepersonell. I denne oppgaven er kun de best dokumenterte smerteevalueringverktøyene inkludert. Verktøyenes bruksevne og pålitelighet undersøkes hovedsakelig ved analysemetoder som undersøker validitet og reliabilitet. Validiteten betraktes som det viktigste målet og gir et svar på om scoringsverktøyet måler det som er ment å måle, mens reliabiliteten viser om man kan stole på målingene som er gjort (32, 33).

Forskning frem til nå viser ingen klare grenser for når man skal benytte de ulike verktøyene, men de fleste hevder at egenrapportering av smerte hvert fall ikke kan benyttes før 3 års alder. Hvilke verktøy som kan benyttes avhenger av en rekke faktorer som kognitiv funksjon, kulturell bakgrunn, og generell utvikling. Verktøy basert på ansiktsuttrykk og «rainbow pain scale» anses som enklere å bruke enn skalaer basert på tallverdier som f.eks. VAS, NRS og CAS. Observasjonsverktøyene inkluderer mange av de samme variablene. Ansiktsuttrykk, gråt/verbaltuttrykk og kroppsposisjon er tre variabler som går igjen som gode indikatorer for smerte. Det er usikkert om disse verktøyene kan benyttes ved langvarige og kroniske smerter, da smerteuttrykket trolig endrer karakter med tiden.

Det er gjort mye forskning på smertescoringsverktøy, men artiklene avdekker også kunnskapshull. Det trengs videre forskning og mer konkrete anbefalinger for når man skal benytte de ulike verktøyene, både mht. barnas alder og for ulike type smerter. Vekting av de ulike elementene innenfor observasjonsskalaene er også et spørsmål for fremtiden. Bør alle variabler vektet likt, eller bør det være et mer komplekst og presist vektingssystem? Det trengs utvikling og validering av observasjonsverktøy for mer langvarige og kroniske smerter hos barn. Kanskje ville det være bedre å kombinere observasjon og egenrapporteringsverktøy til dette formålet? Et annet tema som er gjennomgående er behovet for bedre vurdering av konteksten rundt smerten, dvs. fysiske, kulturelle og sosiale aspekter.

Frem til nå er det vist at smertescoringsverktøy er lite brukt i praksis. En viktig årsak til dette kan være at det er mange verktøy å velge i, samtidig som det mangler klare retningslinjer for når man bør benytte de ulike. Videre forskning kan bidra til klarere retningslinjer for bruk av verktøyene. Det er også behov for enklere og raskere måte å evaluere smerte på. Fremtidige undersøkelser bør fortsette å undersøke mer følsomme og spesifikke mål for smerteintensiteten til barn.

Siden smerte også påvirkes av viktige emosjonelle komponenter er det viktig at foreldre/familie/omsorgspersoner er med på å trygge og trøste barnet for å begrense negative emosjoner. I tillegg kan de bidra med kunnskap rundt barnets personlighet og tidligere smerteerfaring for å optimalisere kartlegging av barnets smerte og behandling (55).

9. Litteratur

1. Legeforening. TN. Smerte hos barn og unge 2002-30-10 [updated 2002-30-10. Available from: <https://tidsskriftet.no/2002/10/bokoversikt/smerte-hos-barn-og-unge>.
2. Soetenga D, Frank J, Pellino TA. Assessment of the validity and reliability of the University of Wisconsin Children's Hospital Pain scale for Preverbal and Nonverbal Children. *Pediatric nursing*. 1999;25(6):670-6.
3. Strout TD, Baumann MR. Reliability and validity of the modified preverbal, early verbal pediatric pain scale in emergency department pediatric patients. *International emergency nursing*. 2011;19(4):178-85.
4. Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. *Pain*. 2001;93(2):173-83.
5. Donna Lee Wong CMB. Pain in children: Comparison of assesment scales *Pediatric nursing*. 1988;14/No.1.
6. von Baeyer CL, Jaaniste T, Vo HLT, Brunsdon G, Lao HC, Champion GD. Systematic Review of Self-Report Measures of Pain Intensity in 3- and 4-Year-Old Children: Bridging a Period of Rapid Cognitive Development. *The journal of pain : official journal of the American Pain Society*. 2017;18(9):1017-26.
7. Saul R, Peters J, Bruce E. Assessing acute and chronic pain in children and young people. *Nursing standard (Royal College of Nursing (Great Britain) : 1987)*. 2016;31(10):51-63.
8. UpToDate. Evaluation and management of pain in children 2010-mai [updated 2019-03-01. Available from: <https://www.uptodate.com/contents/evaluation-and-management-of-pain-in-children>.
9. Helsebibloteket. Symptomer og tilstander - Smerte 2010-04-22 [updated 2015-03-26. Available from: <https://www.helsebiblioteket.no/retningslinjer/palliasjon/4.symptomer-og-tilstander/smerte/definisjon>.
10. Manworren RC, Stinson J. Pediatric Pain Measurement, Assessment, and Evaluation. *Seminars in pediatric neurology*. 2016;23(3):189-200.
11. McGrath PA. Evaluating a child's pain. *Journal of pain and symptom management*. 1989;4(4):198-214.
12. Beltramini A, Milojevic K, Pateron D. Pain Assessment in Newborns, Infants, and Children. *Pediatric annals*. 2017;46(10):e387-e95.
13. Peden V, Choonara I, Vater M. Validating the Derbyshire Children's Hospital Pain Tool in children aged 6-12 years. *Journal of child health care : for professionals working with children in the hospital and community*. 2005;9(1):59-71.
14. von Baeyer CL, Spagrud LJ. Systematic review of observational (behavioral) measures of pain for children and adolescents aged 3 to 18 years. *Pain*. 2007;127(1):140-50.
15. Abu-Saad HH, Pool H, Tulkens B. Further validity testing of the Abu-Saad Paediatric Pain Assessment Tool. *Journal of advanced nursing*. 1994;19(6):1063-71.
16. von Baeyer CL. Children's self-report of pain intensity: what we know, where we are headed. *Pain research & management*. 2009;14(1):39-45.
17. Huguet A, Stinson JN, McGrath PJ. Measurement of self-reported pain intensity in children and adolescents. *Journal of psychosomatic research*. 2010;68(4):329-36.

18. Chang J, Versloot J, Fashler SR, McCrystal KN, Craig KD. Pain assessment in children: validity of facial expression items in observational pain scales. *The Clinical journal of pain*. 2015;31(3):189-97.
19. Tomlinson D, von Baeyer CL, Stinson JN, Sung L. A systematic review of faces scales for the self-report of pain intensity in children. *Pediatrics*. 2010;126(5):e1168-98.
20. Bulloch B, Garcia-Filion P, Notricia D, Bryson M, McConahay T. Reliability of the color analog scale: repeatability of scores in traumatic and nontraumatic injuries. *Acad Emerg Med*. 2009;16(5):465-9.
21. Bird J. Selection of pain measurement tools. *Nursing standard (Royal College of Nursing (Great Britain))* : 1987). 2003;18(13):33-9.
22. Hesselgard K, Larsson S, Romner B, Stromblad LG, Reinstrup P. Validity and reliability of the Behavioural Observational Pain Scale for postoperative pain measurement in children 1-7 years of age. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies*. 2007;8(2):102-8.
23. Fanciullo GJ, Cravero JP, Mudge BO, McHugo GJ, Baird JC. Development of a new computer method to assess children's pain. *Pain medicine (Malden, Mass)*. 2007;8 Suppl 3:S121-8.
24. Gulur P, Rodi SW, Washington TA, Cravero JP, Fanciullo GJ, McHugo GJ, et al. Computer Face Scale for measuring pediatric pain and mood. *The journal of pain : official journal of the American Pain Society*. 2009;10(2):173-9.
25. de la Vega R, Roset R, Castarlenas E, Sanchez-Rodriguez E, Sole E, Miro J. Development and testing of painometer: a smartphone app to assess pain intensity. *The journal of pain : official journal of the American Pain Society*. 2014;15(10):1001-7.
26. Sun T, West N, Ansermino JM, Montgomery CJ, Myers D, Dunsmuir D, et al. A smartphone version of the Faces Pain Scale-Revised and the Color Analog Scale for postoperative pain assessment in children. *Paediatric anaesthesia*. 2015;25(12):1264-73.
27. Bringuier S, Picot MC, Dadure C, Rochette A, Raux O, Boulhais M, et al. A prospective comparison of post-surgical behavioral pain scales in preschoolers highlighting the risk of false evaluations. *Pain*. 2009;145(1-2):60-8.
28. Alves MM, Carvalho PR, Wagner MB, Castoldi A, Becker MM, Silva CC. Cross-validation of the Children's and Infants' Postoperative Pain Scale in Brazilian Children. *Pain practice : the official journal of World Institute of Pain*. 2008;8(3):171-6.
29. Stewart B, Lancaster G, Lawson J, Williams K, Daly J. Validation of the Alder Hey Triage Pain Score. *Archives of disease in childhood*. 2004;89(7):625-30.
30. Hamill JK, Lyndon M, Liley A, Hill AG. Where it hurts: a systematic review of pain-location tools for children. *Pain*. 2014;155(5):851-8.
31. Beyer JE, Wells N. The assessment of pain in children. *Pediatr Clin North Am*. 1989;36(4):837-54.
32. Ellingsen SDoS. Forståelse av kvantitativ helseforskning - en introduksjon og oversikt *Nordisk Tidsskrift for Helseforskning*. 2009;2.
33. Geri LoBiondo-Wood JH. *Nursing Research: Methods and Critical Appraisal for Evidence-Based Practice Nursing Research: Methods and Critical Appraisal for Evidence-Based Practice* St. Louis: Missouri Elsevier 2014. p. 289-310.
34. Beyer JE, Knapp TR. Methodological issues in the measurement of children's pain. *Children's health care : journal of the Association for the Care of Children's Health*. 1986;14(4):233-41.
35. Pripp AH. Validitet *Tidsskrifte norske legeforening*. 2018 Utgave 13

36. Glen S. Statistics How To 2015 [cited 2020 23.01.2020]. Available from: <https://www.statisticshowto.datasciencecentral.com/convergent-validity/>.
37. Ginty AT. Psychometric Properties New York, NY: Springer New York; 2013 [1563-4]. Available from: https://doi.org/10.1007/978-1-4419-1005-9_480.
38. Richard Fellows AL. Research Methods for construction United Kingdom: Wiley-Blackwell Publishing 2008. 295 p.
39. Braut GS. Statistisk signifikans: Store norske leksikon 2018 [updated 26. juni 2018]. Available from: [https://snl.no/statistisk signifikans](https://snl.no/statistisk_signifikans).
40. English FW. The SAGE Guide to Educational Leadership and Management Brown LI, editor: SAGE publications 2015. 584 p.
41. Langvik E. Testbruk og misbruk - kvalitetskrav til testene og de som benytter dem. : Utdanning.no; 2016 [updated 23. februar 2017]. Available from: <https://veilederforum.no/content/testbruk-og-misbruk-%E2%80%93-kvalitetskrav-til-testene-og-de-som-benytter-dem>.
42. Hadden KL, von Baeyer CL. Global and specific behavioral measures of pain in children with cerebral palsy. The Clinical journal of pain. 2005;21(2):140-6.
43. Abu-Saad HH, Kroonen E, Halfens R. On the development of a multidimensional Dutch pain assessment tool for children. Pain. 1990;43(2):249-56.
44. Andersen RD, Langius-Eklof A, Nakstad B, Bernklev T, Jylli L. The measurement properties of pediatric observational pain scales: A systematic review of reviews. Int J Nurs Stud. 2017;73:93-101.
45. Bai J, Jiang N. Where Are We: A Systematic Evaluation of the Psychometric Properties of Pain Assessment Scales for Use in Chinese Children. Pain management nursing : official journal of the American Society of Pain Management Nurses. 2015;16(4):617-31.
46. Bailey B, Daoust R, Doyon-Trottier E, Dauphin-Pierre S, Gravel J. Validation and properties of the verbal numeric scale in children with acute pain. Pain. 2010;149(2):216-21.
47. Beyer JE, Denyes MJ, Villarruel AM. The creation, validation, and continuing development of the Oucher: a measure of pain intensity in children. Journal of pediatric nursing. 1992;7(5):335-46.
48. Beyer JE, Knott CB. Construct validity estimation for the African-American and Hispanic versions of the Oucher Scale. Journal of pediatric nursing. 1998;13(1):20-31.
49. Beyer JE, Turner SB, Jones L, Young L, Onikul R, Bohaty B. The alternate forms reliability of the Oucher pain scale. Pain management nursing : official journal of the American Society of Pain Management Nurses. 2005;6(1):10-7.
50. Bieri D, Reeve RA, Champion GD, Addicoat L, Ziegler JB. The Faces Pain Scale for the self-assessment of the severity of pain experienced by children: development, initial validation, and preliminary investigation for ratio scale properties. Pain. 1990;41(2):139-50.
51. Bosenberg A, Thomas J, Lopez T, Kokinsky E, Larsson LE. Validation of a six-graded faces scale for evaluation of postoperative pain in children. Paediatric anaesthesia. 2003;13(8):708-13.
52. Castarlenas E, Miro J, Sanchez-Rodriguez E. Is the verbal numerical rating scale a valid tool for assessing pain intensity in children below 8 years of age? The journal of pain : official journal of the American Pain Society. 2013;14(3):297-304.
53. Cohen LL, Lemanek K, Blount RL, Dahlquist LM, Lim CS, Palermo TM, et al. Evidence-based assessment of pediatric pain. J Pediatr Psychol. 2008;33(9):939-55; discussion 56-7.

54. Crellin D, Sullivan TP, Babl FE, O'Sullivan R, Hutchinson A. Analysis of the validation of existing behavioral pain and distress scales for use in the procedural setting. *Paediatric anaesthesia*. 2007;17(8):720-33.
55. Dubois A, Bringuier S, Capdevilla X, Pry R. Vocal and verbal expression of postoperative pain in preschoolers. *Pain management nursing : official journal of the American Society of Pain Management Nurses*. 2008;9(4):160-5, 5.e1.
56. Fournier-Charriere E, Tourniaire B, Carbajal R, Cimerman P, Lassaunge F, Ricard C, et al. EVENDOL, a new behavioral pain scale for children ages 0 to 7 years in the emergency department: design and validation. *Pain*. 2012;153(8):1573-82.
57. Franck LS, Greenberg CS, Stevens B. Pain assessment in infants and children. *Pediatr Clin North Am*. 2000;47(3):487-512.
58. Garra G, Singer AJ, Domingo A, Thode HC, Jr. The Wong-Baker pain FACES scale measures pain, not fear. *Pediatric emergency care*. 2013;29(1):17-20.
59. Gharaibeh M, Abu-Saad H. Cultural validation of pediatric pain assessment tools: Jordanian perspective. *Journal of transcultural nursing : official journal of the Transcultural Nursing Society*. 2002;13(1):12-8.
60. Gomez RJ, Barrowman N, Elia S, Manias E, Royle J, Harrison D. Establishing intra- and inter-rater agreement of the Face, Legs, Activity, Cry, Consolability scale for evaluating pain in toddlers during immunization. *Pain Res Manag*. 2013;18(6):e124-8.
61. Hunter M, McDowell L, Hennessy R, Cassey J. An evaluation of the Faces Pain Scale with young children. *Journal of pain and symptom management*. 2000;20(2):122-9.
62. Johnson EK, McQueen JM, Huettig F. Toddlers' language-mediated visual search: they need not have the words for it. *Q J Exp Psychol (Hove)*. 2011;64(9):1672-82.
63. Keck JF, Gerkenmeyer JE, Joyce BA, Schade JG. Reliability and validity of the Faces and Word Descriptor Scales to measure procedural pain. *Journal of pediatric nursing*. 1996;11(6):368-74.
64. Kochman A, Howell J, Sheridan M, Kou M, Shelton Ryan EE, Lee S, et al. Reliability of the Faces, Legs, Activity, Cry, and Consolability Scale in Assessing Acute Pain in the Pediatric Emergency Department. *Pediatric emergency care*. 2017;33(1):14-7.
65. Luffy R, Grove SK. Examining the validity, reliability, and preference of three pediatric pain measurement tools in African-American children. *Pediatric nursing*. 2003;29(1):54-9.
66. Mahon P, Holsti L, Siden H, Strahlendorf C, Turnham L, Giaschi D. Using colors to assess pain in toddlers: validation of "the rainbow pain scale"-a proof-of-principle study. *Journal of pediatric oncology nursing : official journal of the Association of Pediatric Oncology Nurses*. 2015;32(1):40-6.
67. Miro J, Castarlenas E, Huguet A. Evidence for the use of a numerical rating scale to assess the intensity of pediatric pain. *European journal of pain (London, England)*. 2009;13(10):1089-95.
68. Miro J, Huguet A. Evaluation of reliability, validity, and preference for a pediatric pain intensity scale: the Catalan version of the faces pain scale--revised. *Pain*. 2004;111(1-2):59-64.
69. Page MG, Katz J, Stinson J, Isaac L, Martin-Pichora AL, Campbell F. Validation of the numerical rating scale for pain intensity and unpleasantness in pediatric acute postoperative pain: sensitivity to change over time. *The journal of pain : official journal of the American Pain Society*. 2012;13(4):359-69.

70. Quinn BL, Sheldon LK, Cooley ME. Pediatric pain assessment by drawn faces scales: a review. *Pain management nursing : official journal of the American Society of Pain Management Nurses*. 2014;15(4):909-18.
71. Stinson JN, Kavanagh T, Yamada J, Gill N, Stevens B. Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. *Pain*. 2006;125(1-2):143-57.
72. Tsze DS, von Baeyer CL, Bulloch B, Dayan PS. Validation of self-report pain scales in children. *Pediatrics*. 2013;132(4):e971-9.
73. van Dijk M, Koot HM, Saad HH, Tibboel D, Passchier J. Observational visual analog scale in pediatric pain assessment: useful tool or good riddance? *The Clinical journal of pain*. 2002;18(5):310-6.
74. Villarruel AM, Denyes MJ. Pain assessment in children: theoretical and empirical validity. *ANS Advances in nursing science*. 1991;14(2):32-41.
75. Voepel-Lewis T, Merkel S, Tait AR, Trzcinka A, Malviya S. The reliability and validity of the Face, Legs, Activity, Cry, Consolability observational tool as a measure of pain in children with cognitive impairment. *Anesthesia and analgesia*. 2002;95(5):1224-9, table of contents.
76. von Baeyer CL, Spagrud LJ, McCormick JC, Choo E, Neville K, Connelly MA. Three new datasets supporting use of the Numerical Rating Scale (NRS-11) for children's self-reports of pain intensity. *Pain*. 2009;143(3):223-7.
77. Anja Hetland Smeland KS, Hanne Reinertsen, Eva Carlsen. Smertevurdering av barn 0-18 år Helsebibloteket.no: Oslo universitetssykehus 2017 [updated 22.05.2017]. Available from:
<https://www.helsebibloteket.no/fagprosedyrer/ferdige/smertevurdering-av-barn#preparation>.
78. Beyer JE, McGrath PJ, Berde CB. Discordance between self-report and behavioral pain measures in children aged 3-7 years after surgery. *Journal of pain and symptom management*. 1990;5(6):350-6.
79. Carl L. von Baeyer SJF, Elizabeth A. Stanfordc, Mark Watsond, Christine T. Chambers. Response biases in preschool children's ratings of pain in hypothetical situations. *European Journal of Pain* 13 (2009) 209-13.
80. Hillstrom B. 2002 [updated 2014]. Available from:
<http://www.amandasrainbow.com/childrens-rainbows/>.
81. McCarffery MR, MS, FAAN. Choosing a faces pain scale. *Nursing* 2002;32(5):68.
82. Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S. The FLACC: a behavioral scale for scoring postoperative pain in young children. *Pediatric nursing*. 1997;23(3):293-7.
83. Peden V, Vater M, Choonara I. Validating the Derbyshire Children's Hospital Pain Tool: a pilot study. *Paediatric anaesthesia*. 2003;13(2):109-13.
84. Schultz AA, Murphy E, Morton J, Stempel A, Messenger-Rioux C, Bennett K. Preverbal, Early Verbal Pediatric Pain Scale (PEPPS): development and early psychometric testing. *Journal of pediatric nursing*. 1999;14(1):19-27.
85. Shields BJ, Cohen DM, Harbeck-Weber C, Powers JD, Smith GA. Pediatric pain measurement using a visual analogue scale: a comparison of two teaching methods. *Clin Pediatr (Phila)*. 2003;42(3):227-34.
86. Shields BJ, Palermo TM, Powers JD, Grewe SD, Smith GA. Predictors of a child's ability to use a visual analogue scale. *Child Care Health Dev*. 2003;29(4):281-90.

87. Suraseranivongse S, Santawat U, Kraiprasit K, Petcharatana S, Prakkamodom S, Muntraporn N. Cross-validation of a composite pain scale for preschool children within 24 hours of surgery. *Br J Anaesth.* 2001;87(3):400-5.
88. Tarbell SE, Cohen IT, Marsh JL. The Toddler-Preschooler Postoperative Pain Scale: an observational scale for measuring postoperative pain in children aged 1-5. Preliminary report. *Pain.* 1992;50(3):273-80.
89. von Baeyer CL. Children's self-reports of pain intensity: scale selection, limitations and interpretation. *Pain research & management.* 2006;11(3):157-62.
90. von Baeyer CL. Numerical rating scale for self-report of pain intensity in children and adolescents: recent progress and further questions. *European journal of pain (London, England).* 2009;13(10):1005-7.
91. Faiz KW. VAS- visuell analog skala *Tidsskrifte norske legeförening* 2014;Utgave 3

10. Appendix

Søkestreng i PubMed

Search (((("child"[MeSH Terms] OR "child"[All Fields]) OR ("child"[MeSH Terms] OR "child"[All Fields] OR "children"[All Fields]) OR preschool[All Fields] OR ("pediatrics"[MeSH Terms] OR "pediatrics"[All Fields] OR "pediatric"[All Fields])) AND (("pain measurement"[MeSH Terms] OR ("pain"[All Fields] AND "measurement"[All Fields]) OR "pain measurement"[All Fields] OR ("pain"[All Fields] AND "assessment"[All Fields]) OR "pain assessment"[All Fields]) OR ("pain measurement"[MeSH Terms] OR ("pain"[All Fields] AND "measurement"[All Fields]) OR "pain measurement"[All Fields]) OR (("pain"[MeSH Terms] OR "pain"[All Fields]) AND ("weights and measures"[MeSH Terms] OR ("weights"[All Fields] AND "measures"[All Fields]) OR "weights and measures"[All Fields] OR "scale"[All Fields])) OR (("pain"[MeSH Terms] OR "pain"[All Fields]) AND tool[All Fields]) OR (("pain"[MeSH Terms] OR "pain"[All Fields]) AND tools[All Fields]) OR "comfort scale"[All Fields] OR "NRS"[All Fields] OR FLACC[All Fields] OR "PPMP"[All Fields] OR "PPPM"[All Fields] OR "CHEOPS"[All Fields] OR "OPS"[All Fields] OR "POCIS"[All Fields] OR "CHIPPS"[All Fields])) AND (validation[All Fields] OR validate[All Fields] OR ("psychometrics"[MeSH Terms] OR "psychometrics"[All Fields] OR "psychometric"[All Fields]) OR reliability[All Fields] OR reliable[All Fields]) AND ("humans"[MeSH Terms] AND (English[lang] OR Norwegian[lang] OR Swedish[lang])) AND ("child"[MeSH Terms:noexp] OR "infant"[MeSH Terms:noexp] OR "infant"[MeSH Terms] OR "child, preschool"[MeSH Terms]))