



LUNDS UNIVERSITET
Medicinska fakulteten

Smärtskattningsskalor till för tidigt födda barn

En systematisk litteraturstudie

Pain assessment scales adapted for preterm infants

A systematic review

Författare: Ingela Noredal Throbäck

Handledare: Marlene Malmström

Magisteruppsats

Våren 2019

Lunds universitet
Medicinska fakulteten
Programnämnden för omvårdnad, radiografi samt reproduktiv, perinatal och sexuell hälsa
Box 157, 221 00 LUND

Lunds universitet
Medicinska fakulteten
Programnämnden för omvårdnad, radiografi samt reproduktiv, perinatal och sexuell hälsa
Box 157, 221 00 LUND

Smärtskattningsskalor till för tidigt födda barn

En systematisk litteraturstudie

Pain assessment scales adapted for preterm infants

A systemic review

Författare: Ingela Noredal Throbäck

Handledare: Marlene Malmström

Magisteruppsats

Våren 2019

Abstrakt

Bakgrund: Obehandlad smärta hos för tidigt födda barn kan leda till både omedelbara och långtgående negativa effekter. För att kunna ge barnen en adekvat smärtbehandling är det viktigt att regelbunden smärtskatta dem, men då de inte kan kommunicera verbalt är de beroende av att andra kan tolka deras smärtsignaler. **Syfte:** Att identifiera smärtskattningsskalor riktade mot för tidigt födda barn samt undersöka deras validitet och reliabilitet. **Metod:** En systematisk litteraturstudie med arton inkluderade artiklar. **Resultat:** Fjorton smärtskattningsskalor anpassade till för tidigt födda barn identifierades. Alla visade god eller mycket god reliabilitet. Gällande validitet skiljde det sig åt mellan de olika instrumenten. **Slutsats:** Det går inte att rekommendera en specifik smärtskattningsskala baserat på denna litteraturstudie. Det är därför viktigt att fortsätta undersöka de mest lovande i större, kliniska studier och säkerställa deras validitet, reliabilitet, klinisk användbarhet och relevans.

Nyckelord

För tidigt födda barn, smärtskattning, smärtskattningsskalor, reliabilitet, validitet

Innehållsförteckning

Problemområde	6
Bakgrund	7
Perspektiv och utgångspunkter	7
Neonatal omvårdnad	8
Smärta och smärtlindring hos för tidigt födda barn	9
Smärtskattning av för tidigt födda barn	10
Syfte	12
Metod	12
Urval	13
Datainsamling	13
Analys av data.....	14
Statistiska metoder som används i de inkluderade artiklarna	15
Forskningsetiska avvägningar	15
Resultat.....	16
Smärtskattningsskalor	16
Bevis över validitet och reliabilitet	17
ALPS-neo	17
BIIP	17
BPSN	18
COMFORTneo	18
COVERS	19
CPSPN	19
CRIES	20
NFCS-R	20
NIAPAS.....	20
NIPS	21
N-PASS	22
PASPI	23
PAT.....	23
PIPP-R	24
Diskussion	25
Metoddiskussion	25
Resultatdiskussion.....	27
Smärtskattningsskalornas roll i patient- och familjecentrerad vård.....	29

Konklusion och implikationer.....	30
Referenser.....	31
Tabell 1. Sökschema för sökning i PubMed, Cinahl och SweMed+.....	38
Figur 1. Flödesschema över litteratursökningen	40
Tabell 2. Presentation av artiklarna som ingår i litteraturstudien.....	41
Tabell 3. Översikt över de fjorton smärtskattningsskalor som identifierats i studien.	50

Problemområde

Barn som är födda innan graviditetsvecka 37 definieras som för tidigt födda (WHO, 2018). När dessa barn föds behöver de ibland hjälp för att överleva eftersom deras organ, främst hjärna och lungor, inte är färdigutvecklade. De vårdas ofta på speciella intensivvårdsavdelningar för nyfödda och för tidigt födda barn (neonatal-IVA). Under sin tid på neonatal-IVA utsätts de ofta för nödvändiga men smärtsamma ingrepp som blodprovstagningar och intubationer. Det är viktigt med adekvat smärtbehandling eftersom det har visat sig att obehandlad smärta, förutom de omedelbara fysiologiska förändringarna som hypoxemi, hypertension och bradykardi även kan leda till mer långtgående effekter som märks först när barnet är äldre såsom sämre motoriska- och kognitiva förmågor samt beteendeförändringar (Mangat, Oei, Chen, Quah-Smith & Schmölder, 2018).

Trots att det är så viktigt med en tillräcklig smärtlindring av de för tidigt födda barnen finns det idag inte någon egentlig konsensus om vilka metoder som är mest lämpliga att använda för att smärtskatta dem. Smärtskattningen kan antingen göras med hjälp av observationer av eventuella beteendeförändringar hos barnet och/eller med fysiologiska mätmetoder (Andersen, Munsters, Vederhus, & Gradin, 2017). En studie med patienter från neonatal-IVA i 18 europeiska länder visade att endast en tredjedel av barnen blev bedömda med hjälp av en smärtskattningsskala under vårdtiden och endast 10 % blev bedömda dagligen (Anand et al., 2017). Enligt en norsk-svensk studie är det i Sverige vanligt att smärtskattningsskalor används på neonatal-IVA men att sjuksköterskorna inte litar på och känner sig bekväma med att använda dem (Andersen et al., 2017). Om personalen inte litar på att instrumentet mäter det som det avser att mäta kan detta innebära en ökad risk för att de för tidigt födda barnen inte blir korrekt smärtskattade och får adekvat smärtlindring. Samma problem kan uppkomma om ens smärtskattningsskala inte på ett adekvat sätt klarar av att uppmärksamma smärta hos barnen. Inadekvat smärtskattning riskerar således att resultera i att de för tidigt födda barnen i onödan utsätts för obehandlad smärta vilket kan påverka dem i akutskedet men också senare i livet varför utökad kunskap om smärtskattning hos för tidigt födda barn är av stor betydelse.

Bakgrund

Perspektiv och utgångspunkter

Enligt FN:s barnkonvention (UNICEF, 2018) har ett barn rättigheter både som en egen individ och som del av sin familj, och för att tillgodose detta på bästa sätt bör vården runt det för tidigt födda barnet vara person- och familjecentrerad (Institute for patient- and family-centered care, 2019). Personcentrerad vård innebär att barnet inte får bli sin diagnos, ”prematuren” utan ska istället ses som en unik person med egna resurser och behov (Svensk sjuksköterskeförening, 2016). Det som är speciellt med barn, och då särskilt yngre barn som saknar förmåga att kommunicera verbalt, är att familjen blir extra viktig. Familjecentrerad vård innebär att barnets föräldrar inkluderas och uppmuntras till att delta i barnets omvårdnad, samarbete och informationsutbyte mellan föräldrar och vårdpersonal prioriteras, barnet bör vårdas tillsammans med sina föräldrar om detta är möjligt, och att man även uppmärksammar föräldrarnas behov av hjälp (Socialstyrelsen, 2014). Genom att föräldrarna är ständigt närvarande i barnets liv även under sjukhusvistelsen hjälper man barnet och föräldrarna att knyta an till varandra vilket är viktigt för barnets utveckling och välmående men även för att föräldrarna ska känna sig bekväma i sin föräldraroll (Socialstyrelsen, 2014).

Forskning har visat att genom att involvera föräldrarna tidigt i barnens vård kan man motverka långtidseffekterna av att födas för tidigt (Jackson & Wigert, 2015). Föräldrarna är de som är konstanta för barnet och genom att de lär känna sitt barn och dess signaler kan de hjälpa till att förbättra omvårdnaden. De lär sig känna igen ansiktsuttryck som signalerar smärta och särskilja mellan olika typer av gråt och skrik. En studie av Khin Hla et al. (2014) visade att föräldrar ofta smärtskattade högre än sjuksköterskor, och att detta kan innebära att barn får bättre smärtlindring om föräldrarna hjälper till med att smärtskattningen.

Dessutom genom att involvera föräldrarna i barnens smärtlindring kan föräldrarnas stress minskas avsevärt eftersom forskning har visat att en av de saker som stressar föräldrar mest är att se sina barn påverkade av smärta (Jackson & Wigert, 2015).

Neonatal omvårdnad

Miljön i livmodern är anpassad för att skapa en optimal miljö för det ömtåliga fostret med näring och värme, fostervatten som hjälper till att dämpa omgivande ljud och ljus (Jackson & Wigert, 2015). När barnet föds för tidigt är dess hjärna inte mogen nog för att ta emot alla de stimuli, t.ex. starkt ljus, höga ljud, störd sömn, smärtsamma procedurer och mycket aktivitet runt om, som den utsätts för. Hjärnan börjar bildas redan under 5:e-6:e graviditetsveckan (gestationsålder), och i den 10:e graviditetsveckan börjar nervcellerna föröka sig väldigt snabbt och det bildas upp till 200 000 nervceller/min (Lagercrantz, 2007; 2015). Nervcellerna bildas i cellager i kanten på hjärnans hålrum (ventriklarna) och migrerar via gliatrådar sedan till hjärnbarken. I graviditetsvecka 20–25 börjar nervcellerna bilda det neurala nätverket genom att bilda dendritter och axoner. Det är under denna fas som vissa nervceller fortsätter förgrena sig medan andra dör, dvs. den programmerade hjärncellsdöden (Lagercrantz, 2007; 2015). Detta styrs från början genom gener men under den sista trimestern kontrolleras det främst av yttre sensoriska stimuli och det är detta som avgör hur den individuella hjärnan utvecklas. Eftersom hjärnan hos för tidigt födda barn inte är tillräckligt utvecklad klarar den inte av multipla stimuli då den inte kan avgränsa eller sortera intrycken vilket är viktigt att beakta då man vårdar dessa barn (Lagercrantz, 2015).

Enligt WHO (2018) räknas barn som föds innan graviditetsvecka 37 som för tidigt född. Det finns även underkategorier som baseras på i vilken graviditetsvecka barnet föds, extremt för tidigt födda föds innan vecka 28, mycket för tidigt innan vecka 32 och måttligt för tidigt innan vecka 37. Newborn individualized developmental care and assessment program (NIDCAP) är en omvårdnadsmodell som är utvecklad för att vårda för tidigt födda barn (Als, 1986). Det centrala för modellen är att skydda den omogna hjärnan och att inse att även det för tidigt födda barnet har erfarenheter och självreglerande strategier som den förmedlar till sin omgivning (Kleberg, 2013). Barnen vårdas i en miljö som ska vara lugn med dämpat ljus och ljud, vården planeras så att de får möjlighet till perioder av ostörd sömn, man stödjer barnets självreglerande försök till välbefinnande, ger det möjligheter att ligga hud-mot-hud med sina föräldrar och ger stöd till barnet så att det hittar bekväma kroppsställningar som t.ex. fosterställning med armar och ben nära kroppen och händerna nära ansiktet (Jackson & Wigert, 2015). En annan viktig del i NIDCAP är familjecentrerad vård genom att tydliggöra att föräldrarna är det viktigaste i barnets liv och se till så att de blir delaktiga i barnets vård så mycket och tidigt som möjligt (Kleberg, 2013).

Smärta och smärtlindring hos för tidigt födda barn

Inte förrän i slutet på 1980-talet började man inom vården rutinmässigt behandla smärta hos för tidigt födda barn, eftersom det tidigare ansågs att de inte hade förmågan att känna smärta (Hall & Anand, 2014). Senare forskning har visat att även för tidigt födda barn har utvecklade nociceptorer och sensoriska känsel fibrer, men att den utvecklade hjärnan har svårt att avgränsa var smärtan kommer ifrån och även har sämre hämmande mekanismer (Vinall & Grunau, 2014). Förutom den omedelbara responsen på smärta har det också påvisats att barnen utvecklar ett smärtminne och därför reagerar kraftigare vid smärtsamma upplevelser senare i livet som t.ex. vaccinationer vid 3–4 månaders ålder (Mangat et al., 2018). Om de utsätts för upprepad smärta kan det skapas ett hypersensitivt tillstånd där barnen får svårt att särskilja mellan smärta och beröring, och det kan dröja kvar länge (Norman & Fellman, 2015). För tidigt födda barn vårdade på neonatal-IVA utsätts ofta för över hundra smärtsamma ingrepp under sina första veckor i livet (Mangat et al., 2018; Vinall & Grunau, 2014).

När det gäller smärtlindring till för tidigt födda barnen bör smärtan i första hand förebyggas genom att t.ex. provtagningar utförs när barnet är mätt, när det ammar eller tröstas, samt att samla alla provtagningar till ett fåtal tillfällen. Dock är det viktigt att inte utföra andra omvårdnadsåtgärder i samband med provtagningarna för att på så sätt minimera stimuli (Lagercrantz, 2007). Läkemedel som t.ex. opioider används ofta för att behandla smärta hos de för tidigt födda barnen men de kan ha allvarliga biverkningar såsom andningsdepression, hypotension och försämrad tillväxt och utveckling av hjärnan (Anand et al., 2018). För att minimera dessa risker använder man sig i stor grad också av icke-farmakologiska metoder som att barnet ligger i en bra position gärna hud-mot hud (kängurumetoden) med någon av sina föräldrar, att det är nymatat under det smärtsamma momentet, samt att det får sukros/glukos-lösningar (Mangat et al., 2018). Ett exempel på en bra kroppsställning är s.k. ”facilitated tucking”, då barnet med hjälp av en förälders eller vårdpersonals hand ligger på sidan i fosterställning. Detta ger barnet både stabilitet och hudkontakt, vilket har påvisats minska smärta och stress (Jackson & Wigert, 2013). Perorala sockerlösningar som sukros och glukos anses också dämpa smärtan även om man inte vet hur eller varför detta fungerar (Stevens, Craig, Johnston, Harrison & Ohlsson, 2011). Även för tidigt födda barn har ett sugbehov och att ge barnen en tröstnapp, gärna tillsammans med peroral sockerlösning, kan också minska smärtan (Lindh & Lundqvist, 2013). Det är denna multimodala strategi där

icke-farmakologisk och farmakologisk behandling kombineras parallellt med regelbunden smärtskattning som idag utgör smärtbehandlingen för barn födda för tidigt (Socialstyrelsen, 2014).

Smärtskattning av för tidigt födda barn

För att säkerställa att en person inte har obehandlad akut eller kronisk smärta bör sjuksköterskor regelbundet kartlägga eventuell smärta hos sina patienter. Det är denna systematiska beskrivning som sedan ligger till grund för patientens smärtbehandling (Almås, Valand, Bilicz & Berntzen, 2002). Det finns tre olika sätt att mäta smärta på: 1) patienterna rapporterar själv sin smärta, 2) observation av förändringar i deras beteende, och 3) fysiologiska förändringar som t.ex. variationer i hjärtfrekvens eller halten kortisol (Andersen, Langius-Eklöf, Nakstad, Bernklev & Jylli, 2017). Beteendeförändringar som man tittar på hos de för tidigt födda barnen är ansiktsuttryck (öppen mun, sammanpressade stängda ögon, rynkade ögonbryn och en djup näs-läppfåra), sugförmåga, bröstkorgrörelser, vakenhet och tröstbarhet (Lindh & Lundqvist, 2013). Att smärtskatta för tidigt födda barn innebär särskilda utmaningar för sjuksköterskan eftersom de varken kan uttrycka sig verbalt eller är tillräckligt kognitivt utvecklade för att självrapportera sin smärta, istället är de beroende av att andra har förmågan att uppfatta och tolka deras tecken på smärta (Manworren & Stinson, 2016). Dessutom visar för tidigt födda barn mindre ansikts- och kroppsrörelser vid smärta jämfört med fullgångna barn vilket ytterligare försvårar smärtskattningen (Larsson & Frostell, 2016).

För att undvika både över- och underbehandling av smärta samt för att utvärdera resultatet av eventuell smärtlindring bör man regelbundet smärtskatta de för tidigt födda barnen (Andersen et al., 2017). Enligt internationella rekommendationer bör detta göras var 4-6:e timme (Anand et al., 2017). Enligt Socialstyrelsen (2014) bör alla neonatal-IVA ha tydliga rutiner för regelbunden smärtskattning och den bör ske med hjälp av validerade smärtskattningsskalor som är anpassade till barnets mognadsgrad och ålder, samt till den specifika situationen. Då för tidigt födda barn behöver hjälp av människor i sin omgivning för att förmedla deras signaler, har speciella smärtskattningsskalor tagits fram som är baserade på att barnens smärta manifesterar sig som fysiologiska och/eller beteendeförändringar (Andersen et al., 2017). Ett problem är att många smärtskattningsskalor som används till tidigt födda barn är anpassade för akut smärta i samband med procedurer som t.ex. provtagning och inte för mer pågående smärta (Anand et al., 2017). Barn vårdade på neonatal-IVA utsätts ofta för pågående smärta

t.ex. i samband med ventilatorbehandling eller postoperativt. Denna typ av smärta leder ofta till att barnen får en förhöjd basal smärtnivå vilket innebär att då de utsätts för ytterligare smärta är de uttryckslösa, får försämrad sömn och matspjälkning, och man kan se en minskad variabilitet i fysiologiska parametrar (Norman & Fellman, 2015).

Det finns en mängd olika smärtskattningsskalor beskrivna i litteraturen som är anpassade till barn födda för tidigt, t.ex. Premature infant pain profile-revised (PIPP-R; Gibbins et al., 2014), och Astrid Lindgren children's hospital pain scale for preterm and sick newborn Infants (ALPS-Neo; Lundqvist et al., 2014) PIPP-R används vid procedursmärta och består av sju delar: barnets gestationsålder, hur upprört barnet är, förändringar i dess hjärtfrekvens och saturation under det smärtsamma stimuli, beteenden som rynkade ögonbryn, ihopknipna ögon och fördjupning av näs-läppfåran (Gibbins et al., 2014). ALPS-Neo är tänkt att användas för pågående smärta och baseras på observationer av fem olika beteenden hos barnet: ansiktsuttryck, andningsmönster, extremiteternas tonus, hand/fot aktivitet och nivån på barnets aktivitet dvs. lugn eller rastlös (Lundqvist et al., 2014). Några av smärtskattningsskalorna används oftare än andra men det finns ingen konsensus varken nationellt eller internationellt om vilken eller vilka som är lämpligast att använda eller om någon är bättre under särskilda omständigheter (Andersen et al., 2018). Olika varianter av ALPS är den smärtskattningsskala som används på flest av de svenska neonatalavdelningarna samt neonatal-IVA (Andersen et al., 2018).

Obehandlad smärta hos för tidigt födda barn kan få långtgående effekter på både deras motoriska och kognitiva förmågor. De kan dessutom utveckla smärteminnen som leder till en försämrad smärttolerans senare i livet. Det finns många olika smärtskattningsskalor utvecklade för att använda på för tidigt födda barn, men det finns i nuläget ingen konsensus för vilka skalor som är mest lämpliga att använda. Det är då upp till de olika avdelningarna att välja en smärtskattningsskala som de ska använda med risk för att de för tidigt födda barnen inte får en tillräcklig smärtlindring utan riskerar att förutom den akuta smärtan också drabbas av framtida sekvele. Författaren vill med denna studie kartlägga existerande smärtskattningsskalor, samt deras validitet, som blivit framtagna för att specifikt utvärdera smärta, både procedursmärta och pågående smärta, hos för tidigt födda barn.

Syfte

Syftet med denna systematiska litteraturstudie är att beskriva vilka smärtskattningsskalor som är utvecklade för att användas i omvårdnadsarbetet med för tidigt födda barn och om/hur dessa är testade avseende validitet och reliabilitet.

Metod

Studien genomfördes som en systematisk litteraturstudie. Enligt Statens beredning för medicinsk och social utvärdering (SBU; 2018) ska en sådan innehålla: ett väl utformat syfte, tydligt formulerade villkor för litteratursökningen och selektion av artiklar, de artiklar som hittats efter sökningen ska relevansgranskas, kvalitetsbedömas, och till slut bedömdes det sammanlagda resultatet från dem med hjälp av en narrativ sammanställning.

Förutom att undersöka vilka smärtskattningsskalor som fanns tillgängliga för att användas till för tidigt födda barn, var syftet också inriktad på om/hur smärtskattningsskalorna var testade psykometrisk, dvs. hur de var testade angående validitet och reliabilitet. Validitet är ett mått på hur väl ett instrument mäter det man vill att det ska mäta, medan reliabilitet är ett mått på en metods tillförlitlighet (Polit & Beck, 2017). För att reliabiliteten ska vara hög måste resultatet vara oberoende av vem som gör mätningen (inter-rater reliabilitet) samt ge samma resultat vid upprepade mätningar (intra-rater reliabilitet). Ett annat sätt att mäta reliabilitet är att mäta den interna konsistensen vilket är ett mått på om de olika indikatorerna i smärtskattningsskalorna mäter samma sak dvs. de för tidigt födda barnens smärtnivå.

Validitet kan delas upp i tre huvudsakliga komponenter, nämligen innehålls-, kriterie- och begreppsvaliditet. Innehållsvaliditet mäter hur väl innehållet i smärtskattningsskalorna täcker området, dvs. om skalans innehåll är bra är sannolikheten att det når det uppsatta målet större och att barnets smärta mäts/skattas på ett adekvat sätt. Kriterievaliditet innebär att man jämför smärtskattningsskalan med en annan validerad smärtskattningsskala som optimalt ska anses vara en "golden standard". Begreppsvaliditet innebär att smärtskalan verkligen mäter det som den är avsedd att göra, dvs. barnens smärta. För att ha en bra begreppsvaliditet ska smärtskattningsskalan klara av att särskilja mellan när barnet upplever smärta, stress eller ingenting alls. Diskriminativ validitet ingår i begreppsvaliditeten och testar om de olika

indikatorerna i smärtskattningsskalan är oberoende av varandra. Detta innebär att alla indikatorer individuellt ska kunna särskilja mellan smärta, stress eller ingenting alls (Polit & Beck, 2017).

Urval

Inklusionskriterierna var artiklar som var publicerade från år 2000 och framåt, och de skulle vara skrivna på svenska, engelska, norska eller danska. Studierna i artiklarna skulle vara kvantitativa och kunde vara experimentella (Randomized controlled trials) eller kvasi-experimentella. För att inkluderas i arbetet skulle artiklarna vara publicerade och "peer-review"-granskade (dvs. artiklarna har blivit granskade innan publicering av en grupp forskare kunniga inom området). Smärtskattningsskalorna skulle vara baserade på enbart beteendeförändringar eller kombinerade med fysiologiska parametrar. De för tidigt födda barnen kunde vara vårdade på både neonatalavdelning (Internationell level II vilket innebär att sjukhuset har både obstetriska och pediatrika avdelningar) och neonatal-IVA (Level III, tillgång till neonatal-IVA; Finnström et al., 1997), samtliga barn som vårdades på dessa avdelningar ingick, och smärtskattningsskalorna skulle vara specifikt anpassade för barn födda för tidigt. Artiklar som var systematiska reviews, hade en kvalitativ ansats, undersökte fullgångna spädbarn eller var anpassade till specifika länders kontext exkluderades.

Datainsamling

För att få en överblick i hur forskningsläget såg ut samt för att identifiera lämpliga sökord gjordes först en pilotsökning i databasen PubMed. Litteratursökning genomfördes sedan med hjälp av följande databaser: Pubmed, Cinahl och SweMed+. Anledningen till att dessa databaser användes var för att få en så bred täckning som möjligt. Pubmed innehåller artiklar om medicin och omvårdnad, Cinahl är specialiserad på omvårdnadsforskning och SweMed+ innehåller vårdforskning på svenska, norska, danska och engelska (Forsberg & Wengström, 2015). Sökorden som användes var premature, infant, preterm, infant, neonatal prematurity, pain, assessment, scale och measurement. Sökstrategin anpassades för vardera databas genom att använda MeSH-termer i PubMed och SweMed+, och Cinahl Headings i Cinahl (se Tabell. 1). Artiklarnas abstrakt lästes igenom, och om de svarade mot inklusions- och exklusionskriterierna granskades de i fulltext. Sedan genomfördes en sekundär sökning genom att referenslistorna till de artiklar som hittades i primärsökningen gick igenom manuellt. Detta gjordes för att hitta eventuella artiklar som inte hittades vid primärsökningen.

Litteratursökningen genererade 263 potentiella artiklar (143 i PubMed, 120 i Cinahl och 0 i SweMed+), se Tabell. 1, varav 202 exkluderades efter att titlarna lästs igenom. Av de kvarvarande 61 artiklarna var arton dubletter och kunde tas bort. Abstrakt lästes igenom hos de kvarvarandes 43 artiklarna och ytterligare 13 exkluderades då de inte uppfyllde inklusionskriterierna. De kvarvarande 30 lästes igenom till fullo och 9 exkluderades eftersom de undersökte fullgångna spädbarn (n=2), skalorna var inte baserade på beteendeförändringar (n=2), de utvärderade inte smärtskattningsskalor (n=1), eller hade inte genomgått en etisk granskning (n=4). Referenslistorna på de kvarvarande 21 artiklarna granskades och ytterligare tre potentiella artiklar identifierades. De genomgick samma bedömningsprocess som övriga artiklar. En av artiklarna exkluderades efter att abstract lästs igenom, och en efter att hela artikeln lästs igenom då den inte genomgått någon etisk granskning. Litteratursökningsprocessen sammanfattas i Figur 1.

För att kontrollera de 22 studiernas relevans användes mallen för bedömning av relevans från SBU (2018). Utvärderingen av artiklarna och beslut om slutgiltig inkludering i studien gjordes ensamt av författaren. De utvalda artiklarna (n=20) granskades med hjälp av Critical Appraisal Skills Programme (CASP, 2019) för kvalitetsgranskning av kohort-studier. Endast sektion A användes eftersom det är här som artiklarnas studiedesign utvärderas. Den bestod av åtta punkter, och svaren kunde vara ja, nej eller oklart. De två första frågorna anses vara screening-frågor och avser hur väl syftet är beskrivet samt hur urvalet skett. Om svaret på någon av dessa blev nej exkluderades artikeln. Bekvämlighets- och konsekutivt urval ansågs vara lämpligt för litteraturstudien. För att gradera studier gavs ja ett värde av 1, nej och oklart värdet 0, och det totala värdet för en artikel blev 8 (Tripathi & Dursing, 2015). Artiklarna ansåg ha hög kvalitet om de fick 7–8 poäng, medelkvalitet vid 5–6 och låg kvalitet vid 1–4.

Analys av data

En statistisk metaanalys var inte möjlig att göra eftersom de inkluderade artiklarna skiljde sig åt avseende urvalsgrupper, metoder och hur man presenterade resultatet. Istället sammanfattades resultaten från de olika artiklarna i en narrativ och beskrivande form fokuserande deras användningsområde och psykometriska utfall. De 18 artiklarna som klarade kvalitetsgranskningen presenteras i en artikelöversikt (se Tabell. 2) och det som ingår i summeringen är: 1) författare, publiceringsår, titel och ursprungsland, 2),

smärtskattningsskalans namn 3) design och urval, 4) inklusions- och exklusionskriterier, 5) validitet, 6) reliabilitet och 7) artikelkvalitet.

Statistiska metoder som används i de inkluderade artiklarna

I artiklarna bestäms reliabiliteten med hjälp av intern konsistens, inter-rater och intra-rater. Dessa beräknas med hjälp av intraclass korrelation (ICC), Cohens kappa (κ) och Cronbachs alfa (α). ICC tar hänsyn både till korrelationen inom smärtskattningsskalan samt likhet mellan instrumentets domäner, och ICC $<0,5$ är dåligt, $0,5-0,75$ måttligt, $0,75-0,90$ god och $>0,90$ mycket god (Koo & Li, 2016). Cohens κ anses vara lämplig eftersom den tar hänsyn till slumpen. Riktlinjer för värdering är: $\kappa <0$ ingen överensstämmelse, $0-0,20$ dåligt, $0,21-0,40$ tveksam, $0,41-0,60$ måttlig, $0,61-0,80$ god och $0,81-1,00$ mycket god (Landis & Kock, 1977). Det finns inga direkta riktlinjer på värden för Cronbachs α men ett förslag är: $\alpha <0,5$ dåligt, $0,51-0,60$ svag, $0,61-0,7$ tveksam, $0,71-0,80$ acceptabel, $0,81-0,90$ god och $0,91-1,00$ mycket god (Tavakol & Dennick, 2011).

Variabiliteten uttrycks i artiklarna antingen med hjälp av ett spridningsdiagram (covariance matrix method) eller med hjälp av en korrelationskoefficient (Spearman's rho (ρ), Pearsons korrelationskoefficient, Pearsons eller Spearman's rangkorrelation). Korrelationen kan variera mellan $-1,00$ (perfekt negativ korrelation) via 0 (ingen korrelation) till $1,00$ (perfekt positiv korrelation). Det finns inga direkta riktlinjer på hur man avgör vilka värden korrelationen ska ha för att anses vara stark utan detta avgörs från fall till fall (Polit & Beck, 2017). T-test (parametriska värden) och Wilcoxon signed-tank test, Kruskal Wallis test, ANOVA och Friedmans test (icke-parametriska värden) användes för att upptäcka skillnader i medelvärde mellan smärta och icke-smärta, eller smärta och stress (Polit & Beck, 2017).

Forskningsetiska avvägningar

Enligt Forsberg och Wengström (2015) är det viktigt att göra etiska överväganden gällande urval och presentation av resultatet innan man påbörjar en systematisk litteraturstudie. Därför är alla inkluderade artiklar i denna studie granskade och godkända av en etisk kommitté. Alla artiklar som ingått i sökningen är redovisade, och inklusions- samt exklusionskriterierna tydligt beskrivna för att visa att ingen selektionsbias förelåg (Forsberg & Wengström, 2015). Eftersom enbart oidentifierat material angående deltagarna i de analyserade artiklarna används utgör den ingen risk för någon enskild person.

Resultat

Arton artiklar inkluderades i studien och de var publicerade mellan 2004 och 2019 (se Tabell. 2). De flesta studierna ägde rum på level III care-avdelningar men några skedde även på level II care. Ursprungsländerna var Australien (n=2), Finland (n=1), Indien (n=1), Kina (n=1), Kanada (n=4), Nederländerna (n=2), Schweiz (n=1), Sverige (n=1), Taiwan (n=1), Thailand (n=1) och USA (n=3). Antalet deltagare i studierna varierade från 12 till 202 (m=63). Sjutton av studierna var prospektiva kohortstudier där man gjorde mätningar på en grupp för tidigt födda barn medan en studie (Lundqvist et al., 2014) var en klinisk observationsstudie där man inte utsatte deltagarna för någon speciell behandling. Fyra studier (Desai, Nanavati, Jasani & Kabra, 2017; Hummel, Puchalski, Creech, & Weiss, 2008; Lundqvist et al., 2014; McNair, Ballantyne, Dionne, Stephens and Stevens, 2004) var inriktade på smärtskattningsskalor för pågående smärta medan övriga på procedursmärta. I alla studierna smärtskattade man för tidigt födda barn, från 24 veckors gestationsålder till fullgångna barn. I sex av studierna (Cignacco, Mueller, Hamers & Gessler, 2004; McNair et al., 2004; Caljouw et al., 2007; Gibbins et al., 2014; Huang et al., 2018; Woragidpoonpol, Tiansawad, Mesukko & Klunklin, 2018) delade man upp de för tidigt födda barnen i olika åldersgrupper för att testa om smärtskattningsskalorna kunde upptäcka skillnader korrelerade med gestationsålder.

Smärtskattningsskalor

Smärtskattningsskalorna som identifierades i de arton inkluderade artiklarna var ALPS-neo, BIIP, BPSN, CPSPN, COMFORTneo, COVERS, CRIES, NFCS-R, NIAPAS, NIPS, N-PASS, PASPI, PAT och PIPP-R. I Tabell 3 presenteras de med namn, om de var anpassade för procedur- eller pågående smärta, deras bedömningspunkter och poäng, samt fysiologiska och beteendeindikatorer. Två var endimensionella (ALPS-neo och NFCS-R), dvs. innehöll enbart beteendemässiga indikatorer medan övriga var multidimensionella, dvs innehöll både fysiologiska och beteendemässiga indikatorer. Alla fjorton innehöll en värdering av ansiktsuttryck där man utgick ifrån det klassiska smärtansiktet med rynkade ögonbryn, stängda sammanpressade ögon, djup näs-läppfåra, öppen mun och en synlig skålformad tunga. När det gäller fysiologiska indikatorer var hjärtfrekvens den vanligaste indikatorn (n=11) följd av SaO₂ (n=7). Nio av smärtskattningsskalorna användes för att enbart skatta procedursmärta (BIIP, BPSN, CPSPN, COVERS, NFCS-R, NIAPAS, NIPS, PASPI och PAT), tre för att enbart skatta pågående smärta (ALPS-neo, COMFORTneo och CRIES) och två för att skatta både procedur- och pågående smärta (N-PASS och PIPP-R).

Bevis över validitet och reliabilitet

Alla fjorton smärtskattningsskalorna var undersökta angående validitet och reliabilitet på något sätt. Det vanligaste sättet att testa validitet var att jämföra smärtskattningsskalan med en annan tidigare validerad smärtskattningsskala (kriterievaliditet), och det var gjort med alla utom två (ALPS-neo och CRIES). De vanligaste skalorna att jämföra med var PIPP/PIPP-R (n=5) och NIPS (n=2). Det förekom också att man jämförde med skalor där en sjuksköterska eller förälder uppskattade barnets smärta med hjälp av visual analog scale (VAS; n=2) och numerisk skala (NRS; n=1). En annan metod för att bekräfta validitet var att se om skalorna uppvisade begreppsvaliditet. Det gjordes för alla smärtskattningsskalor förutom tre (ALPS-neo, CRIES, och PAT). En studie (Kappesser et al., 2019) undersökte diskriminativ validitet för BPSN, NFCS-R, NIPS, N-PASS och PIPP-R. Skalornas reliabilitet testades genom att två eller fler användares resultat för samma mätning jämfördes (inter-rater reliabilitet). Dessutom kontrollerade man intern konsistens. Det var bara CRIES som inte var testad för interrater reliabilitet och intern konsistens. Endast tre skalor (BPSN, NIAPAS och N-PASS) testades för intra-rater reliabilitet.

ALPS-neo

ALPS-neo är endimensionell smärtskattningsskala, baserad på fem indikatorer, som är framtagen för att användas vid pågående smärta. I en observationsstudie av Lundqvist et al. (2014) undersöktes instrumentets innehållsvaliditet genom att det utvärderades både med en intern (25 sjuksköterskor och en neonatolog) och en extern expertgrupp (tre disputerade sjuksköterskor) i ett första steg, och av fem disputerade sjuksköterskor som studerade videosekvenser i ett andra skede. Reliabiliteten för instrumentet utvärderades i två steg (n=125 och n=96, resp.) med mycket god inter-rater reliabilitet (ICC=0,91) och intern konsistens ($\alpha=0,95$).

BIIP

BIIP är ett multidimensionellt instrument som baseras på åtta indikatorer. Dess validitet och reliabilitet har undersökts i två studier (Holsti & Grunau, 2007; Holsti, Grunau, Oberlander & Osiovič, 2008) med vardera 100 bedömningar av videoinspelade barn. Kriterievaliditeten när BIIP jämfördes med NIPS var endast måttlig ($k=0,64$; Holsti & Grunau, 2007).

Begreppsvaliditeten för instrumentet var god och man kunde se en signifikant ökning från

baslinjen ($1,0 \pm 1,8$) vid en smärtsam procedur (blodprovstagning; $5,3 \pm 2,6$, CL -4,8 till -3,7, $p < 0,001$) och en sänkning vid återhämtningsfasen ($1,8 \pm 0,3$, CL 2,8 till 4,1, $p < 0,0001$; Holsti & Grunau, 2007). Holsti et al. (2008) jämförde BIIP-värden under tre faser vid en blodprovstagning, ett blöjbyte samt blodprovstagning efter ett blöjbyte. Detta jämfördes med förändringar i hjärtfrekvens. Signifikanta förändringar i BIIP och hjärtfrekvens från baslinjen kunde ses i samband med stressmomenten i alla tre procedurerna. Paired t-test visade att BIIP var signifikant högre under blodprovstagningen jämfört med blöjbytet ($5,1$ vs $3,1$, $t=5,1$, $p < 0,0001$). De barn som fick blöjorna bytta innan blodprovstagningen hade ännu högre BIIP-värden ($5,1$ vs $6,0$, $t=2,3$, $p < 0,03$). Inter-rater reliabilitet ($ICC=0,80-0,92$) och intern konsistens ($\alpha=0,82$) var god för instrumentet (Holsti & Grunau, 2007).

BPSN

BPSN är en multidimensionell smärtskattningsskala som baseras på nio indikatorer (tre fysiologiska och sex beteenden). Den har testats i två studier med 288 (Cignacco et al., 2004) respektive 42 videoinspelade bedömningar (Kappesser et al., 2019) vid procedursmärta. Kriterievaliditet var måttlig till mycket god och fastställdes genom jämförelser med PIPP ($\alpha=0,907$), VAS ($\alpha=0,855$; Cignacco et al., 2004), NFCS-R ($k=0,864$), NIPS ($k=0,794$), N-PASS ($k=0,834$) och PIPP-R ($k=0,828$; Kappesser et al., 2019). Begreppsvaliditet fastställdes eftersom instrumentet kunde separera olika faser i en smärtsam procedur (blodprovstagning; $F=41,27$; Cignacco et al., 2004). Inter-rater reliabilitet var mycket god i båda studierna ($\alpha=0,9792$ och $ICC=0,992$). Kappesser et al. (2019) undersökte även intern konsistens ($\alpha=0,843$) och Cignacco et al. (2004) intra-rater reliabilitet ($\alpha=0,976-0,991$).

COMFORTneo

COMFORTneo är en multidimensionell skala baserad på en fysiologisk och sju beteendeindikatorer som är testad för procedursmärta. I en studie av Caljouw et al. (2007) undersökte man instrumentet med hjälp av 114 bed-side bedömningar av en smärtsam procedur (kapillär provtagning i hälen). Kriterievaliditet var god i jämfördes med en sjuksköterskas uppskattning av smärtan i form av VAS. Inter-rater reliabilitet ($\kappa=0,85$ innan proceduren och $0,93$ efter) och intern konsistens ($\alpha=0,76$ innan och $0,86$ efter) var god.

COVERS

COVERS utvecklades för att kunna smärtskatta alla nyfödda barn oavsett i vilken gestationsvecka de var födda i och oberoende av deras sjukdomstillstånd (Hand, Noble, Geiss, Wozniak, & Hall, 2010). Den är multidimensionell och bygger på sex indikatorer, två fysiologiska och fyra beteendeindikatorer. I en studie av Hand et al. (2010) undersöktes dess validitet med ett urval på 21 barn. Kriterievaliditet var god ($k=0,84$) när en enda observatör smärtskattade barnen vid tre tillfällen under en procedur (blöjbyte eller kapillärprovtagning i hälen): innan (baslinje), under och efter (baslinje), och jämförde COVERS med PIPP. För att undersöka skalans begreppsvaliditet jämförde man värdena mellan en kapillär provtagning med ett blöjbyte. Den klarade av att separera de olika faserna i båda procedurerna (dess poäng ökade signifikant från baslinjen), och mellan den kapillära provtagningen (medel 7,3) och blöjbytet (medel 4,9). I en studie av O'Sullivan, Rowley, Ellis, Faasse och Petrie (2016) användes COVERS vid bed-side smärtskattning av 80 nyfödda barn. Kriterievaliditet fastställdes genom resultaten från COVERS jämfördes med både PAT ($r=0,81$) och moderns smärtskattning av barnet (NRS, $r=0,74$). Begreppsvaliditet undersöktes genom att man jämförde en smärtsam procedur (kapillär provtagning i hälen) med en icke-smärtsam (blöjbyte; $z=-7,22$). Inter-rater reliabiliteten ($ICC=0,801-0,82$) och den interna konsistensen var god ($\alpha=0,74-0,78$).

CPSPN

CPSPN är en multidimensionell smärtskattningsskala som bygger på tre beteendeindikatorer och en fysiologisk. Den har analyserats i en studie där man hade ett urval på 19 för tidigt födda barn som var planerade att genomgå en smärtsam procedur (Woragidpoonpol et al., 2018). På dessa barn gjordes totalt 53 bedömningar av två varandra oberoende bedömare. Instrumentets kriterievaliditet bestämdes genom att den jämfördes med PIPP-R ($r=0,789$), och dess begreppsvaliditet genom att smärtskatta innan, under och efter en kapillär provtagning i hälen ($\chi^2=95,95$). Instrumentets reliabilitet testades på två olika sätt: inter-rater (ICC innan, under och efter var 0,95, 0,91 och 1,00, resp.), och intern konsistens ($k=0,94$) som båda var mycket goda.

CRIES

CRIES är ett multidimensionellt instrument, och bygger på två fysiologiska och tre beteendeindikatorer. som undersökts i samband med pågående, postoperativ smärta. McNair et al. (2004) använde instrumentet i en studie med 52 barn som genomgått en operation. De analyserade instrumentets kriterievaliditeten under de första 72 h postoperativt genom att jämföra med resultatet från PIPP och VAS. Barnen smärtskattades var fjärde timme de första 24 h, och sedan var åttonde timme under nästföljande två dygn samt 2 h efter varje bolusdos med analgetika. Resultaten visade måttlig korrelation mellan instrumenten de första 24 h (ICC 0,55–0,88) men sämre efter längre tid (ICC 0,07–0,26).

NFCS-R

NFCS-R är ett endimensionellt instrument som baseras på det klassiska smärtansiktet hos för tidigt födda barn: rynkade ögonbryn, sammanpressade ögon, näs-läppfåra, stängd mun och synlig tunga (Kappesser et al., 2019). Den har testats i en studie med videoinspelade bedömningar av 42 barn (Kappesser et al., 2019) vid procedursmärta. Instrumentets kriterievaliditet undersöktes genom att det jämfördes med NIPS ($k=0,870$), N-PASS ($k=0,898$), PIPP-R ($k=0,875$) och BPSN ($k=0,864$). Instrumentet hade starkare korrelation med de andra vid högre värden, dvs. ju mer smärta som barnet upplevde. Även diskriminativ validitet testades och alla skalans individuella indikatorer kunde skilja på smärta (blodprovstagning) och stress/icke-smärta (blöjbyte, $p=0,002–0,004$). Inter-rater reliabiliteten (ICC=0,997) var mycket god och intern konsistens ($\alpha=0,983$) god.

NIAPAS

NIAPAS är en multidimensionell smärtskattningsskala speciellt framtagen för att skatta smärta hos för tidigt födda barn (Pölkki, Korhonen, Axelin, Saarela, & Laukkala, 2014; Huang et al., 2018). Den inkluderar fem beteendeindikatorer och tre fysiologiska, och tar även hänsyn till barnets gestationsålder. Dess validitet och reliabilitet har undersökts i två studier (Pölkki et al., 2014; Huang et al., 2018). I studien av Pölkki et al. (2014) gjordes 180 bedömningar på 34 för tidigt födda barn som filmades när de utsattes för antingen en blodprovstagning i hälen eller sugning i deras tracheostomi. Metodens validitet testades genom att barnen skattades innan (1 min), under och efter (1 min) den smärtsamma proceduren (begreppsvaliditet, $F=90,7$). Dessutom värderades NIAPAS genom att man

samtidigt skattade barnen med NIPS (kriterievaliditet; innan $r=0,873$, under $r=0,751$ och efter $r=0,804$, resp.). Eftersom varje smärtskattning gjordes av två olika personer som inte var medvetna om varandras smärtskattningar kunde inter-rater reliabilitet bedömas (innan $r=0,991$, under $r=0,997$ och efter $r=0,995$). Intra-rater reliabilitet analyserades genom att samma smärtskattningar gjordes om efter 2–3 veckor ($r=0,992$ – $1,00$ i de olika faserna). Intern konsistens var måttlig ($\alpha=0,723$). Studien av Huang et al. (2018) baserade på 1080 bedömningar av 90 barn som filmades och smärtskattades i efterhand. Kriterievaliditeten undersöktes via jämförelser med PIPP-R ($r=0,783$) och N-PASS ($r=0,823$).

Begreppsvaliditeten analyserades under både en smärtsam (arteriell blodprovstagning; $x^2=241,152$) och en icke-smärtsam (blöjbyte; $x^2=202,774$) procedur, smärtskattning skedde innan (P0), under (P1) och efter proceduren (P2). Värdena steg från P0 till P1 och sjönk sedan igen vid P2 (smärtsam $1,89 \pm 2,441 - 8,03 \pm 4,301 - 5,37 \pm 4,294$; icke-smärtsam $1,39 \pm 1,396 - 3,93 \pm 3,415 - 2,99 \pm 2,464$). Två olika personer smärtskattade varje barn oberoende av varandra (inter-rater reliabilitet; ICC=0,99). Intern konsistens var god ($\alpha=0,836$). Båda studierna visade att NIAPAS hade bra validitet och reliabilitet.

NIPS

NIPS är ett multidimensionellt instrument som bygger på en fysiologisk indikator och fem beteendeindikatorer som undersökts i tre studier (Holsti & Grunau, 2007; Pölki et al., 2014; Kappesser et al., 2019). För NIPS har kriterievaliditet analyserats genom att instrumentets resultat vid en smärtskattning jämförts med resultatet från NIAPAS ($r=0,751$ – $0,873$; Pölki et al., 2014), BIIP ($r=0,64$, Holsti & Grunau, 2007), NFCS-R ($k=0,759$), N-PASS ($k=0,921$), PIPP-R ($k=0,608$) och BPSN ($k=0,794$; Kappesser et al., 2019). Man undersökte även instrumentets begreppsvaliditet genom att kontrollera hur väl den kunde särskilja mellan tre faser i en kapillär provtagning: innan, vid och efter. En signifikant ökning kunde ses från baslinjen ($2,0 \pm 1,2$) till blodprovstagningen ($4,0 \pm 1,7$) och signifikant sänkning till återhämtningsfasen ($2,3 \pm 1,8$; Holsti & Grunau, 2007). Kappesser et al. (2019) undersökte dessutom instrumentets diskriminativa validitet genom att undersöka varje indikatorns förmåga att skilja på smärta (blodprovtagning) och stress (blöjbyte). Alla indikatorer kunde skilja på smärta och stress förutom ansiktsuttryck ($p=0,058$), och vakenhet ($p=0,088$). Det fanns dock signifikans för att hela instrumentet kunde skilja mellan smärta och stress ($p < 0,001$). Samma studie visade även att reliabilitet i form av inter-rater var mycket god

(ICC=0,993 smärta och 0,973 stress) och intern konsistens god ($\alpha=0,882$ smärta och 0,845 stress).

N-PASS

N-PASS är en av två smärtskattningsskalor som testats på både procedur och pågående smärta. Det är också den enda som även kan användas för att skatta sederingsgrad hos de för tidigt födda barnen. Den är multidimensionell och består av fem indikatorer (gråt, ansiktsuttryck, tonus i extremiteterna, vakenhetsgrad samt förändring i en vital parameter). Tre studier (Hummel et al., 2010, Huang et al., 2018; Kappesser et al., 2019) har undersökt dess validitet och reliabilitet i samband med procedursmärta. Kriterievaliditet undersöktes genom att N-PASS jämfördes med PIPP ($r=0,743$; Hummel et al., 2010), NFCS-R ($k=0,898$), NIPS ($k=0,921$), BPSN ($k=0,834$; Kappesser et al., 2019), PIPP-R ($k=0,894$; Kappesser et al., 2019; $r=0,787$; Huang et al., 2018), och NIAPAS ($r=0,823$; Huang et al., 2018). Två av studierna testade begreppsvaliditet där Huang et al. (2018) smärtskattade de för tidigt födda barnen i samband med både en smärtsam (arteriell blodprovstagning; $\chi^2= 220,681$) och en icke-smärtsam (blöjbyte; $\chi^2= 115,880$) procedur. Smärtskattning skedde innan (P0), vid (P1) och efter (P2) proceduren. Vid både den smärtsamma och den icke-smärtsamma proceduren steg värdet från P0 till P1, och sjönk sedan igen till P2 (smärtsam $0,92 \pm 1,569 - 5,18 \pm 3,111 - 3,19 \pm 3,192$; icke-smärtsam $0,64 \pm 1,226 - 2,19 \pm 2,373 - 1,32 \pm 2,181$). Hummel et al. (2010) jämförde en falsk och en äkta kapillär provtagning i hälen och visade att N-PASS kunde särskilja dessa åt (medel äkta=3,93, medel falsk=0,81, $z=-6,429$). Kappesser et al. (2019) testade diskriminantiv validitet för att fastställa att alla de indikatorer som ingick i N-PASS kunde identifiera smärta, och de visade att alla beteendeindikatorerna klarade detta ($p \leq 0,005$) men inte den fysiologiska ($p=0,014$). Reliabilitet undersöktes i alla tre studierna genom intern konsistens och inter-rater reliabilitet, och smärtskattningsskalan visade goda eller mycket goda resultat i alla studierna (ICC=0,887 och $\alpha=0,86$, Hummel et al., 2010; ICC=0,837 och $\alpha=0,99$; Huang et al. (2018; ICC=0,837 och $\alpha=0,99$; Kappesser et al., 2019). Både Hummel et al. (2010) och Kappesser et al. (2019) kontrollerade om N-PASS kunde upptäcka någon skillnad mellan barn födda vid olika gestationsåldrar men ingen hittade någon signifikant skillnad. Endast Hummel et al. (2010) undersökte intra-rater reliabilitet genom att göra om smärtskattningen vid videoinspelad situation efter en vecka ($r=0,874$) och ett år ($r=0,846$).

Två studier undersökte N-PASS i samband med pågående smärta (Desai et al., 2017; Hummel et al., 2008). Studien av Hummel et al. (2008) baserades på 72 bed-side bedömningar av 46 barn som var intuberade eller vårdades postoperativt, och fick analgetika kontinuerligt och/eller intermittent. Begreppsvaliditet undersöktes genom att smärtskatta innan ($4,86 \pm 3,38$) och efter ($1,81 \pm 1,53$) analgetika. Inter-rater reliabiliteten var mycket god ($ICC=0,965$) och den interna konsistensen god ($\alpha=0,82$). I studien av Desai et al. (2017) undersöktes 42 barn som hade vårdats i invasiv eller non-invasiv ventilator >48 h. I båda studierna jämfördes N-PASS med PIPP, och kriterievaliditeten var måttlig i båda ($r=0,61$ och $r=0,62$, resp.)

PASPI

PASPI är en multidimensionell skala som baseras på två fysiologiska indikatorer och fyra beteendeindikatorer. PASPI konstruerades i en studie av Liaw et al. (2011) med ett urval på 240 bedömningar av 60 för tidigt födda barn. Dess innehållsvaliditet verifierades genom att tio neonatala, kliniska experter utvärderade indikatorerna som ingår i skalan. Dess validitet testades genom att PIPP och VAS användes som smärtskattningsskala jämsides PASPI (kriterievaliditet; $r=0,74-0,83$ och $r=0,75-0,82$, resp.), samt hur bra PASPI kunde separera fyra olika faser (begreppsvaliditet; innan, vid, tre minuter och tio minuter efter) under en kapillär provtagning i hälen ($F=23,72, 56,86, 60,81$ och $96,18$, i resp. fas). Alla bedömningar gjordes av två bedömare oberoende av varandra. PASPI bedömdes ha god till mycket god inter-rater reliabilitet ($ICC=0,88-0,93$) och god intern konsistens ($\alpha=0,84$).

PAT

PAT är en multidimensionell smärtskattningsskala som baseras på nio indikatorer (fyra fysiologiska och fem beteendeindikatorer) och dessutom inkluderas bedömarens uppfattning om barnets smärta. Validering och reliabilitet för PAT undersöktes i två studier (O'Sullivan et al., 2016; Spence, Gillies, Harrison, Johnston, & Nagy, 2005) med urval på 80–144 barn. Instrumentets kriterievaliditet har testats i jämförelse med CRIES (Spence et al., 2005), COVERS (O'Sullivan et al., 2016) och även barnets mammas smärtskattning i form av VAS (Spence et al., 2005). Den hade stark korrelation med CRIES ($r=0,76$) och COVERS ($r=0,81$) men endast svag med moderns smärtskattning ($r=0,38$). Spence et al. (2005) undersökte även inter-rater reliabiliteten som var god ($ICC=0,85$).

PIPP-R

PIPP-R är en omarbetad version av PIPP som är det instrument som tidigare använts mest som smärtskattningsskala (Stevens, Johnston, Petryshen, & Taddio, 1996). Många anser dock att instrumentet är alltför komplicerat för att användas kliniskt. PIPP-R är ett multidimensionellt instrument som baseras på två fysiologiska och tre beteendeindikatorer (Gibbins et al., 2014). Två studier har undersökt instrumentets kriterievaliditet (Huang et al., 2018; Kappesser et al., 2019) genom att jämföra det med N-PASS ($r=0,787$, Huang et al., 2018; $k=0,894$, Kappesser et al., 2019), NIAPAS ($0,783$; Huang et al., 2018), NFCS-R ($k=0,693$), NIPS ($k=0,888$), BPSN ($k=0,828$; Kappesser et al., 2019). Huang et al. (2018) testade begreppsvaliditet genom att barnen smärtskattades under både en smärtsam (arteriell blodprovstagning; $x^2= 251,984$) och en icke-smärtsam (blöjbyte; $x^2= 183,512$) procedur, smärtskattning skedde innan (P0), under (P1) och efter proceduren (P2). Värdena steg från P0 till P1 och sjönk sedan till P2 (smärtsam $0,97 \pm 2,860 - 11,17 \pm 4,903 - 7,81 \pm 5,173$; icke-smärtsam $0,41 \pm 1,72 - 6,83 \pm 4,734 - 4,16 \pm 4,124$). Gibbins et al. (2014) undersökte också begreppsvaliditet genom att kontrollera om PIPP-R kunde särskilja mellan en smärtsam (kapillär blodprovstagning eller sugning i tracheostomin) och en icke-smärtsam (blöjbyte, vändning, temperaturmätning). Resultaten blev signifikant högre under den smärtsamma proceduren ($6,7 \pm 3,0$ vs $4,8 \pm 2,9$). Detta var den enda smärtskattningsskalan som kunde se någon signifikant skillnad mellan barn födda i olika graviditetsveckor med högst värden i den yngsta åldersgruppen för både den smärtsamma proceduren ($p=0,003$) och icke-smärtsamma ($p < 0,001$). Kappesser et al. (2019) undersökte diskriminantiv validitet där indikatorerna för hjärtfrekvens ($p=0,024$), saturation ($p=0,065$) och vakenhetsstadium ($p=0,777$) inte klarade av att skilja på smärta (blodprovstagning) och stress (blöjbyte). Men när man undersökte hela PIPP-R fanns det signifikans för att instrumentet kunde separera smärta från stress ($p < 0,001$). I denna studie kunde det inte ses någon signifikant skillnad mellan barn födda i olika graviditetsveckor ($p=0,505$). Alla tre studierna visade god inter-rater reliabilitet och intern konsistens.

Diskussion

Metoddiskussion

För att få en insikt i hur forskningsläget såg ut samt för att hitta lämpliga sökord gjordes först en pilotsökning. Sökorden anpassades till vardera databasen genom att databasernas ämnesordslista användes. Dessa utprövade sökord användes sedan till studiens primärsökning. Enligt Willmann, Bahtsevani, Nilsson och Sandström (2016) bör en väl utförd litteraturstudie ha både hög sensitivitet och specificitet. Litteratursökningen bakom denna studie verkar ha haft en bra sensitivitet eftersom sökningarna i PubMed och Cinahl genererade många artiklar som var dubletter samt att sekundärsökningen endast genererade tre nya artiklar. Däremot var specificiteten inte hög då sökningarna hittade 263 studier och endast 29 ansågs vara tillräckligt relevanta för att läsas till fullo. Men i detta initiala läget är det viktigare med hög sensitivitet än specificitet (Willmann et al., 2016). För att få en bredare sökning kan man även söka på s.k. grå litteratur, vilket inkluderar forsknings- och utbildningsrapporter samt mer populärvetenskapliga tidskrifter. Detta har inte gjorts i denna studie vilket innebär att en del relevant information kan ha missats. Författaren valde att bara inkludera artiklar publicerade mellan 2000 och 2019 eftersom klinisk forskning kan anses åldras snabbt (Forsberg & Wengström, 2015). Detta innebär att en del artiklar publicerade tidigare inte inkluderats i litteraturstudien. Eftersom endast studier som blivit etiskt granskade och godkända inkluderades exkluderades även några artiklar pga. detta.

De inkluderade artiklarna var genomförda i många olika länder. Det kan innebära att överförbarheten till svensk hälso- och sjukvård kan anses vara begränsad. I denna litteraturstudie bedöms detta dock vara av begränsad betydelse då smärtskattning av för tidigt födda barn bör utföras på ungefär samma sätt oavsett var man befinner sig i världen och att sociala barriärer inte borde påverka så mycket. Det medför också en intressant aspekt med så många olika ursprungsländer eftersom de olika smärtskattningsskalorna ofta är utformade och testade i specifika länder, t.ex. ALPS-neo i Sverige och BIIP i Kanada.

Ett viktigt steg i litteraturstudien var kvalitetsgranskningen. Författaren har tidigare inte gjort kvalitetsgranskningar av artiklar på detta systematiska sätt vilket kan ha påverkat bedömningarna. För att bestämma vilka artiklar som har tillräckligt hög kvalitet för att bli inkluderade i studien kan man efter granskning dela upp dem i tre olika grupper: låg, medel

och hög kvalitet. Artiklar med låg kvalitet bör inte inkluderas (Forsberg & Wengström, 2015). I denna litteraturstudie var det svårt att hitta en lämplig kvalitetsmall eftersom många tillgängliga var mer generellt anpassade för observationsstudier vilket gjorde dem irrelevanta och svåra att använda. Kvalitetsgranskningsmallen som valdes ut för att användas i denna studie (CASP, 2019) var designad för att användas specifikt i kohort-studier och hade tidigare använts i en liknade studie (Tripathi & Dursing, 2015). Granskningsmallen kändes relevant för ändamålet och var samtidigt användarvänlig då den endast bestod av åtta frågor.

Ett problem med många av de inkluderade artiklarna (n=15) var att de använde för små urvalsgrupper. Enligt Anthoine, Moret, Regnault, Sébille och Hardouin (2014) är en lämplig urvalsstorlek 100–200 vilket endast uppnås av två studier (Spence et al., 2005; Gibbins et al., 2014). Men i sex av studierna har barnen bedömts mer än en gång och varje bedömning har räknats som individuell. I tio av artiklarna har det innan studiens start beräknats hur stora urvalsgrupper som behövdes för att säkra slutsatser skulle kunna dras. Detta gjordes med hjälp av poweranalys vilket innebär att man beräknar hur stor urvalsgrupp som behövs för att minimera typ II fel, dvs. falska negativa, och förhindra statistiska felkällor (Polit & Becker, 2017). En annan nackdel med de inkluderade artiklarna var att vissa grupper av för tidigt födda barn är exkluderade från studierna och det därför inte går att dra några generella slutsatser om huruvida smärtskattningsskalorna går att använda på alla för tidigt födda barn. Tre barngrupper som ofta exkluderades var de med neurologiska sjukdomar, medfödda missbildningar samt de som fått muskelavslappnande läkemedel. I endast två av artiklarna (Lundqvist et al., 2014; Woragidpoonpol et al., 2018) exkluderade inga barn baserat på sjukdomstillstånd eller medicinering.

En stor begränsning i denna litteraturstudie är att allt arbete utförts av endast en person. Speciellt arbetet med att relevans- och kvalitetsbedöma de olika artiklarna är sårbart när endast en person utför det. Risken att personliga åsikter, fördomar samt systematiska fel påverkar slutresultatet ökar jämfört med om två eller flera utför och analyserar de olika stegen. Enligt SBU (2018) bör varje steg i en systematisk litteraturstudie utföras av minst två av varandra oberoende granskare. Detta var tyvärr inte möjligt inom ramarna för denna litteraturstudie. För att minimera svagheten kopplat till ensidig relevans- och kvalitetsbedömning har författaren varit noga med att följa de använda gransknings- och bedömningsmallarna.

Resultatdiskussion

Litteraturstudien identifierade fjorton olika smärtskattningsskalor som användes på för tidigt födda barn. Alla visade god eller mycket god inter-rater reliabilitet och intern konsistens. Tre (BPSN, NIAPAS och N-PASS) uppvisade även god intrarater reliabilitet. Validiteten skiljde sig åt mellan de olika smärtskattningsskalorna både med hänseende till hur man testat validitet och resultaten. Den vanligaste smärtskattningsskalan var N-PASS som användes i tre studier med procedurmärta och två med pågående smärta. Instrumentet uppvisade måttlig-god kriterievaliditet jämfört mot ett flertal andra smärtskattningsskalor och bra begreppsvaliditet i tre studier (Hummel et al, 2010; Huang et al., 2018; Kappesser et al., 2019). I studien av Hummel et al. (2008) där den användes vid pågående smärta klarade den av att mäta skillnaden i smärta innan och efter barnen fått analgetika. I studien av Kappesser et al. (2019) påvisades att alla indikatorer i skalan kunde särskilja mellan smärta och stress, vilket också hela skalan kunde. Dock hade bara en av studierna tillräckligt stort urval (Huang et al., 2018) vilket är en nackdel. En annan negativ aspekt var att många patientgrupper exkluderades vilket innebär att det är osäkert om instrumentet kan användas på alla för tidigt födda barn. För att säkerställa att smärtskattningsskalorna verkligen används inom omvårdnaden räcker det inte att de har god validitet och reliabilitet utan vårdpersonalen måste kunna lita på dem och anse att de är användarvänliga. Sjuksköterskor är överlag positiva till att använda smärtskattningsskalor men många anser också att de kan värdera smärtan hos för tidigt födda barn lika bra utan att använda ett instrument (Pölki et al., 2010). I en svensk studie av Nieminen och Bergsten (2015) där vårdpersonalens (n=69) inställning till att använda N-PASS undersöktes uppgav endast 7,2 % att de var negativa till instrumentet. Det som upplevdes som negativt var att instrumentet var svårt att använda eftersom det inte var lämpligt till för tidigt födda barn och att det var svårt att poängsätta denna speciella barngrupp.

Smärtskattningsskalan som används mest i Sverige, ALPS-neo (Andersen et al., 2018) uppvisade god reliabilitet men hade inte blivit testad för varken kriterie-, begrepps- eller diskriminativ validitet. I studien av Lundqvist et al. (2014) testade man enbart innehållsvaliditet, dvs. man utvärderade de inkluderade indikatorerna både externt och intern i två steg med utvärdering emellan. Enligt Polit och Beck (2017) är det viktigt när man rekryterar sina experter att man använder personer med stora kunskaper om ämnet som instrumentet ska mäta och populationen som ska undersökas, och att olika professioner är

representerade, vilket också gjordes i konstruktionen av ALPS-neo. En annan positiv aspekt var att urvalsgruppen bestod av alla barn på neonatal-IVA och inga exkluderades, vilket innebär att skalan går att använda på alla för tidigt födda barn. Dock behöver skalan utvärderas mer utförligt angående begrepps- och diskriminativ validitet.

Det var vanligast att smärtskattningsskalan kombinerade beteendemässiga indikatorer med en eller fler fysiologiska indikatorer. Den vanligaste beteendemässiga indikatorn var ansiktsuttryck som i någon form ingick i alla fjorton smärtskattningsskalor. Några hade bara ansiktsuttryck generellt (n=10) medan andra var mycket mer specifika och utgick från det klassiska smärtansiktet, rynkade ögonbryn, sammanpressade ögon, näsläpp-fåra, stängd mun och synlig tunga, hos både fullgångna och för tidigt födda spädbarn (Lindh och Lundqvist, 2013). Det mest extrema var NFCS-R som endast var baserad på smärtansiktet (Kappesser et al., 2019). Fysiologiska indikatorer anses ibland inte vara specifika nog för att kunna särskilja smärta från stress (Barr, 1998), och studien av Kappesser et al. (2019) stöder till viss del detta påstående. De visade att fysiologiska indikatorer hos vissa smärtskattningsskalor var sämre på att skilja mellan smärta och stress än beteendeindikatorerna. Men de fysiologiska parametrarna kan kanske ändå användas som ett komplement till smärtskattningsskalorna eftersom de ofta finns tillgängliga på övervakningsutrustningen på en neonatal-IVA och därmed är enkla att övervaka och följa.

En viktig aspekt att ta hänsyn till när för tidigt födda barn smärtskattas är att jämfört med fullgångna spädbarn visar de mindre ansikts- och kroppsrörelser (Larsson & Frostell, 2016). För sju av smärtskattningsskalorna (BPSN, COMFORT-neo, CPSPN, CRIES, NIAPAS, N-PASS och PIPP-R) har hänsyn tagits till detta och de för tidigt födda barnen delades upp i grupper baserade på gestationsålder. Endast en av smärtskattningsskalorna, PIPP-R, visade en signifikant skillnad mellan de olika åldersgrupperna (Gibbins et al., 2014) men det är osäkert vad detta innebär. Kanske finns det ingen skillnad i förmågan att visa smärta eller så är de andra smärtskattningsskalorna så bra konstruerade att de eventuella skillnaderna mellan åldersgrupperna inte spelar någon roll. En annan faktor som också kan påverka för tidigt födda barns förmåga att uttrycka smärta är deras tidigare smärtupplevelser (Lindh & Lundqvist, 2013) men detta har inte blivit undersökt för någon smärtskattningsskala. Eftersom för tidigt födda barn ofta spenderar långa perioder på neonatal-IVA och utsätts för många smärtsamma procedurer är detta något som borde undersökas vidare.

Alla studier förutom ALPS-neo hade blivit testade angående kriterievaliditet, och de uppvisade alla god eller mycket god korrelation till andra skalor. Men för att man ska kunna undersöka kriterievaliditet behövs en s.k. ”golden standard”, dvs. en tidigare validerad smärtskattningsskala som man jämför sitt nya instrument med (Polit & Beck, 2017). Problemet är att det inte finns någon ”golden standard” när det gäller smärtskattningsskalor anpassade till för tidigt födda barn. Istället används de andra smärtskattningsskalorna som jämförelse. Så frågan är hur mycket information man egentligen får ut av dessa valideringar eftersom allting hänger på om den jämförande skalan egentligen mäter det den ska. Ett annat sätt att göra jämförelsen är emot skalor där antingen vårdpersonal eller en förälder skattar barnets smärta enligt VAS eller NRS. Dessa är egentligen skalor där patienterna själva ska användas gradera sina smärta. Att som vårdpersonal värdera sin patients smärta på en 10-gradig skala är subjektivt och resultaten måste därför mötas med skepticism. Ett alternativ till en golden standard kan vara att smärtskattningsskalorna istället jämförs med t.ex. utsöndringen av biokemiska stressmarkörer som kortisol eller nya metoder som infra-red spectroscopy (NIRS) och EEG (Lindh & Lundqvist, 2013).

Smärtskattningsskalornas roll i patient- och familjecentrerad vård

Enligt person- och familjecentrerad vård ska föräldrarna till de för tidigt födda barnen uppmuntras till att delta så mycket som de önskar i barnets omvårdnad (Institute for patient- and family-centered care, 2019). En av de situationer som upplevs som extra jobbig för föräldrar är när de ser att deras barn har ont och att de då upplever att de inte har möjlighet eller förmåga att hjälpa dem (Jackson & Wigert, 2015). Om den på avdelningen använda smärtskattningsskalan introduceras för föräldrarna samtidigt som de får information och undervisning i hur den ska användas kan de enkelt bli mer delaktiga i barnets vård. När föräldrar får möjlighet att aktivt delta i smärtbehandling minskar deras stressnivåer avsevärt (De Bernardo, Svelto, Giordano, Sordino & Riccitelli, 2017). En intervjustudie av Palomaa, Korhonen och Pölkki (2016) visade att det absolut viktigaste för föräldrar under tiden på neonatal-IVA var att de fick vara delaktiga i barnets smärtbehandling och på så sätt förbättra sitt barns välbefinnande. De ville få uppmuntran av personalen och känna sig välkomna, och att detta bäst uppnåddes med utförlig information, verbala instruktioner och praktisk visning.

En annan viktig aspekt är att studier har visat att föräldrar och vårdpersonal inte alltid bedömer spädbarns smärta på samma sätt utan föräldrar upplever ofta att vårdpersonalen

undervärderar smärtan (Khin Hla et al., 2014). Om föräldrarna får möjlighet att använda samma smärtskattningsskala som vårdpersonalen använder kan kanske konsensus lättare uppnås angående barnets smärta och smärtlindring vilket kan innebära att föräldrarna får ett ökat förtroende för att vårdpersonalen tar barnets smärta på allvar. För att detta ska vara möjligt måste smärtskattningsskalan inte bara vara psykometriskt testad utan också lättanvänd så att föräldrarna känner att de klarar av det. Den största anledningen till att föräldrar kände sig hindrade att delta i sitt barns vård var att de upplevde att det var för svårt och komplicerat (Palomaa et al, 2016). Om föräldrarna får tillfälle att delta i beslutsfattandet och även får en möjlighet att berätta om sina observationer för personalen, angående t.ex. barnens tecken på smärta, förstärks känslan att vara närvarande för sina barn vilket är en av tankarna bakom familjecentrerad vård (Axelin, Outinen, Lainema, Lehtonen & Franck, 2018).

Konklusion och implikationer

Det är svårt att rekommendera en specifik smärtskattningsskala baserat på resultaten från denna litteraturstudie men N-PASS verkar ha potential. Även ALPS-neo verkar lovande men behöver testas i större studier och verifieras mer angående validitet. De flesta av smärtskattningsskalorna har testats under konstruerade miljöer och ofta har bedömningarna skett via videoinspelningar. Därför är det av yttersta vikt att de testas kliniskt och de under studierna blir en del av det dagliga arbetet på avdelningarna för att deras reliabilitet och validitet samt klinisk användbarhet kan fastställas. Mycket fokus har tidigare lagts på att skapa nya smärtskattningsskalor istället för att vidareutveckla de som redan finns. Denna litteraturstudie har visat att en del av de redan existerande smärtskattningsskalorna har god validitet och reliabilitet, och därmed potential för att kunna användas kliniskt.

I framtiden behöver de bästa smärtskattningsskalorna undersökas bättre angående användbarhet. Om instrumenten inte upplevs som lättanvändbara kommer de aldrig bli en del i det dagliga arbetet, och det är av yttersta vikt att personalen upplever att de faktiskt kan lita på att instrumenten mäter det som de ska. En annan viktig aspekt är att fastställa cut-off-värden för instrumenten så att de kan hjälpa personalen att avgöra när man t.ex. behöver ge analgetika. Genom att säkerställa att den valda smärtskattningsskalan har validitet, reliabilitet och är användarvänlig möjliggörs att vårdpersonal och föräldrar på ett strukturerat och vetenskapligt sätt kan behandla de för tidigt födda barnens smärta och på så sätt förhindra onödigt lidande.

Referenser

Almås, H., Valand, E., Bilicz, J.A., & Berntzen, H. (2002). Smärta. In H. Almås (Ed.), *Klinisk omvårdnad*. (3. uppl., p. 65–114). Stockholm: Liber AB.

Als, H. (1986). Program guide-Newborn Individualized Developmental Care and Assessment Program (NIDCAP): an education and training program for health care professionals. Boston, Copyright, NIDCAP federation International.

Anand, K. J. S., Eriksson, M., Boyle, E. M., Avila-Alvarez, A., Andersen, R. D., Sarafidis, K., ... Carbajal, R. (2017). Assessment of continuous pain in newborns admitted to NICUs in 18 European countries. *Acta Paediatrica*, *106*, 1248–1259.

Andersen, R. D., Langius-Eklöf, A., Nakstad, B., Bernklev, T., & Jylli, L. (2017). The measurement properties of pediatric observational pain scales: a systematic review of reviews. *International journal of nursing studies*, *73*, 93-101.

Andersen, R. D., Munsters, J. A., Vederhus, B. J., & Gradin, M. (2018). Pain assessment practices in Swedish and Norwegian neonatal care units. *Scandinavian journal of caring sciences*, *32*, 1074-1082.

Anthoine, E., Moret, L., Regnault, A., Sébille, V., & Hardouin, J.B. (2014). Sample size used to validate a scale: a review of publications on newly-developed patient reported outcome measures. *Health and quality of life outcomes*, *12*, 1-10.

Axelin, A., Outinen, J., Lainema, K., Lehtonen, L., & Franck, L. (2018). Neonatologists can impede or support parents' participation in decision-making during medical rounds in neonatal intensive care units. *Acta Paediatrica*, *107*, 2100–2108.

Barr, R. (1998). Reflections on measuring pain in infants: dissociation in responsive systems and "honest signaling". *Archives of disease in childhood*, *79*, F152-F156.

*Caljouw, M., Kloos, M., Olivier, M., Heemskerk, I., Pison, W., Stigter, G., & Verhoef, A-M. (2007). Measurement of pain in premature infants with a gestational age between 28 to 37 weeks: validation of the adapted COMFORT scale. *Journal of neonatal nursing*, 13, 13-18.

CASP. (2019). CASP cohort study checklist. Hämtad 2019-03-08 från https://casp-uk.net/wp-content/uploads/2018/01/CASP-Cohort-Study-Checklist_2018.pdf

*Cignacco, E., Mueller, R., Hamers, J. P. H., & Gessler, P. (2004). Pain assessment in the neonate using the Bernese pain scale for neonates. *Early human development*, 78, 125-131.

De Bernardo, G., Svelto, M., Giordano, M., Sordino, D., & Riccitelli, M. (2017). Supporting parents in taking care of their infants admitted to a neonatal intensive care unit: a prospective cohort pilot study. *Italian journal of pediatrics*, 43:36. DOI 10.1186/s13052-017-0352-1

*Desai, S.A., Nanavati, R.N., Jasani, B.B., & Kabra, N. (2017). Comparison of neonatal pain, agitation, and sedation scale with premature infant pain profile for the assessment of acute prolonged pain in neonates on assisted ventilation: a prospective observational study. *Indian journal of palliative care*, 23(3), 287-292.

Finnström, O., Otterblad Olausson, P., Sedin, G., Serenius, F., Svenningsen, N., Thiringer, K., ... Tunell, R. (1997). The Swedish national prospective study on extremely low birthweight (ELBW) infants. Incidence, mortality, morbidity and survival in relation to level of care. *Acta paediatrica*, 86, 503-511.

Forsberg, C., & Wengström, Y. (2015). *Att göra systematiska litteraturstudier: värdering, analys och presentation av omvårdnadsforskning* (4. uppl.). Stockholm: Natur och kultur.

*Gibbins, S., Stevens, B.J., Yamada, J., Dionne, K., Campbell-Yeo, M., Lee, G., Caddell, K., Johnston, C., & Taddio, A. (2014). Validation of the premature infant pain profile-revised (PIPP-R). *Early human development*, 90, 189-193,

Hall, R.W., & Anand, K.J.S. (2014). Pain management in newborns. *Clinics in perinatology*, 41(4), 895-924.

*Hand, I.L., Noble, L., Geiss, D., Wozniak, L., & Hall, C. (2010). COVERS neonatal pain scale: development and validation. *International journal of pediatrics*.

Doi:10.1155/2010/496719

*Holsti, L., & Grunau, R.E. (2007). Initial validation of the behavioral indicators of infant pain (BIIP). *Pain*, 132(3), 264-272.

*Holsti, L., Grunau, R.E., Oberlander, T.F., & Osiovich, H. (2008). Is it painful or not? Discriminant validity of the behavioral indicators of infant pain (BIIP) scale. *Clinical journal of pain*, 24(1), 83-88.

*Huang, X-Z., Li, L., Zhou, J., He, F., Zhong, C-X., & Wang, B. (2018). Evaluation of three pain assessment scales used for ventilated neonates. *Journal of clinical nursing*, 27, 3522-3529.

*Hummel, P., Puchalski, M., Creech, S.D., & Weiss, M.G. (2008). Clinical reliability and validity of the N-PASS: neonatal pain, agitation and sedation scale with prolonged pain. *Journal of perinatology*, 28, 55-60.

*Hummel, P., Lawlor-Klean, P., & Weiss, M.G. (2010). Validity and reliability of the N-PASS assessment tool with acute pain. *Journal of perinatology*, 30, 474-478.

Institute for patient- and family-centered care. (2019). Patient- and family-centered care is working “with” patients and families, rather than just doing “to” or “for” them. Hämtad 2019-05-14 från <http://www.ipfcc.org/about/pfcc.html>

Jackson, K. & Wigert, H. (Red.). (2013). *Familjecentrerad neonatalvård*. Lund: Studentlitteratur AB.

*Kappesser, J., Kamper-Fuhrmann, E., de Laffoile, J., Faas, D., Ehrhardt, H., Franck, L., & Hermann, C. (2019). Pain-specific reactions or indicators of a general stress response? Investigating the discriminant validity of 5 well-established neonatal pain assessment tools. *Clinical journal of pain*, 35(2), 101-110.

Khin Hla, T. Hegarty, M., Russell, P., Drake-Brockman, T.F., Ramgolam, A., & von Ungern-Sternberg, B.S. (2014). Perception of pediatric pain: a comparison of postoperative pain assessments between child, parent, nurse, and independent observer. *Pediatric anaesthesia*, 24(11), 1127–1131.

Kleberg, A. (2013). Familjecentrerad utvecklingsstödande vård enligt NIDCAP. I P. Lundqvist (Red.), *Omvårdnad av det nyfödda barnet*. (2. uppl., s. 41–90). Lund: Studentlitteratur.

Koo, T.K. & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, 15, 155–163.
<http://dx.doi.org/10.1016/j.jcm.2016.02.012>

Lagercrantz, H. (2007). Hjärnans utveckling och mognad. I T. Lindberg, & H. Lagercrantz (Red.), *Barnmedicin* (3. uppl., s. 130–138). Lund: Studentlitteratur.

Lagercrantz, H. (2015). Hjärnans utveckling. I H. Lagercrantz & L. Hellström-Westas & M. Norman (Red.), *Neonatologi*. (2. uppl., s. 269–276). Lund: Studentlitteratur.

Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. DOI:10.2307/2529310

Larsson, B. A., & Frostell, C. (2016). Intensivvård av barn. I A. Larsson & S. Rubertsson (Eds.), *Intensivvård*. (2. uppl., s. 714–727). Stockholm: Liber AB.

*Liaw, J-J., Hang, L., Chou, H-L., Yin, T., Chao, S-C., & Lee, T-Y. (2011). Psychometric analysis of a taiwan-version pain assessment scale for preterm infants. *Journal of clinical nursing*, 21, 89–100.

Lindh, V., & Lundqvist, P. (2013). Smärtbehandling. I P. Lundqvist (Red.), *Omvårdnad av det nyfödda barnet*. (2. uppl., s. 223–247). Lund: Studentlitteratur.

*Lundqvist, P., Kleberg, A., Edberg, A.-K., Larsson, B.A., Hellström-Westas, L., & Norman, E. (2014). Development and psychometric properties of the Swedish ALPS-Neo pain and stress assessment scale for newborn infants. *Acta paediatrica*, 103, 833-839.

Mangat, A. K., Oei, J-L., Chen, K., Quah-Smith, I., & Schmölzer, G. M. (2018). A review of non-pharmalogical treatments for pain management in newborn infants. *Children*, 5(130). doi:10.3390

Manworren, R., & Stinson, J. (2016). Seminars in pediatric neurology pediatric pain measurement, assessment and evaluation. *Seminars in pediatric neurology*, 23(3), 189-200.

*McNair, C., Ballantyne, M., Dionne, K., Stephens, D., & Stevens, B. (2004). Postoperative pain assessment in the neonatal intensive care unit. *Archives of disease in childhood. Fetal and neonatal edition*, 89, F537-541.

Nieminen, C. & Bergsten, P. (2015). *Smärtskattning av för tidigt födda barn på en neonatalavdelning*. (Examensarbete i vårdvetenskap, Uppsala universitet, Sverige). Hämtad från www.diva-portal.org/smash/get/diva2:814558/FULLTEXT01.pdf

Norman, E., & Fellman, V. (2015). Smärta och smärtbehandling. I H. Lagercrantz & L. Hellström-Westas & M. Norman (Red.), *Neonatologi*. (2. uppl., s. 485–494). Lund: Studentlitteratur.

*O’Sullivan, A.T., Rowley, S., Ellis, S., Faasse, K., & Petrie, K.J. (2016). The validity and clinical utility of the COVERS scale and pain assessment tool for assessing pain in neonates admitted to an intensive care unit. *Clinical journal of pain*, 32(1), 51–57.

Palomaa, A-K., Korhonen, A., & Pölkki, T. (2016). Factors influencing parental participation in neonatal pain alleviation. *Journal of pediatric nursing*, 31, 519-527.

Polit, D.F., & Beck, C.T. (2017). *Nursing research. Generating and assessing evidence for nursing practice* (10., uppl.). Philadelphia, PA: Wolters Kluwer.

*Pölkki, T., Korhonen, A., Axelin, A., Saarela, T., & Laukkala, H. (2014). Development and preliminary validation of the neonatal infant pain assessment scale (NIAPAS). *International journal of nursing studies*, 51, 1585-1594.

Pölkki, T., Korhonen, A., Laukkala, H., Saarela, T., Vehviläinen-Julkunen, K., & Pietilä, A-M. (2010). Nurses' attitudes and perceptions of pain assessment in neonatal intensive care. *Scandinavian journal of caring sciences*, 24, 49-55. Doi:10.1111/j1471-6712.2008.00683.x

Socialstyrelsen. (2014). Vård av extremt för tidigt födda barn. Hämtad: 2018-12-17 från <https://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/19503/2014-9-10.pdf>

*Spence, K., Gillies, D., Harrison, D., Johnston, L., & Nagy, S. (2003). A reliable pain assessment tool for clinical assessment in the neonatal intensive care unit. *JOGNN*, 33(5), 80–86.

Statens beredning för medicinsk och social utvärdering. (2018). Vår metod. Hämtad: 2018-11-27 från <http://www.sbu.se/sv/var-metod/>

Stevens, B., Johnston, C., Petryshen, P., & Taddio, A. (1996). Premature infant pain profile: development and initial validation. *Clinical journal of pain*, 12(1), 13-22.

Stevens, B., Craig, K., Johnston, C., Harrison, D., & Ohlsson, A. (2011). Oral sucrose for procedural pain in infants. *The Lancet*, 377(9759), 25–26.

Svensk sjuksköterskeförening. (2016). Personcentrerad vård. Hämtad: 2018-12-16 från [www.swenurse.se/globalassets/01-svensk-sjuksköterskeforening/publikationer-svensk-sjuksköterskeforening/ssf-om-publikationer/svensk_sjuksköterskeforening_om_personcentrerad_vard_oktober_2016.pdf](http://www.swenurse.se/globalassets/01-svensk-sjukskoterskeforening/publikationer-svensk-sjukskoterskeforening/ssf-om-publikationer/svensk_sjukskoterskeforening_om_personcentrerad_vard_oktober_2016.pdf)

Tavakol, M. & Dennick, R. (2011). Making sense of cronbach's alpha. *International journal of medical education*, 2, 53-55. Doi:10.5116/ijme.4dfb.8dfd

Tripathi, T. & Dusing, S.C. (2015). Long-term neurodevelopmental outcomes of infants born late preterm: a systematic review. *Research and reports in neonatology*, 91-111.

UNICEF. (2018). FN:s barnkonvention. Hämtad 2018-12-17 från <https://unicef.se/barnkonventionen/las-texten#full>

Vinall, J., & Grunau, R. E. (2014). Impact of repeated procedural pain-related stress in infants born very preterm. *Pediatric research*, 75(5), 584–587.

Willman, A., Bahtsevani, C., Nilsson, R., & Sandström, B. (2016). *Evidensbaserad omvårdnad. En bro mellan forskning och klinisk verksamhet* (4. uppl.). Lund: Studentlitteratur.

*Woragidpoonpol, P., Tiansawad, S., Mesukko, J., & Klunklin, P. (2018). Development of a clinical pain scale for preterm neonates. *Pacific rim international journal of nursing research*, 22(4), 347-359.

World Health Organization (WHO). (2018). Preterm birth. Hämtad 2019-01-09 från <http://www.who.int/news-room/fact-sheet/detail/preterm-birth>

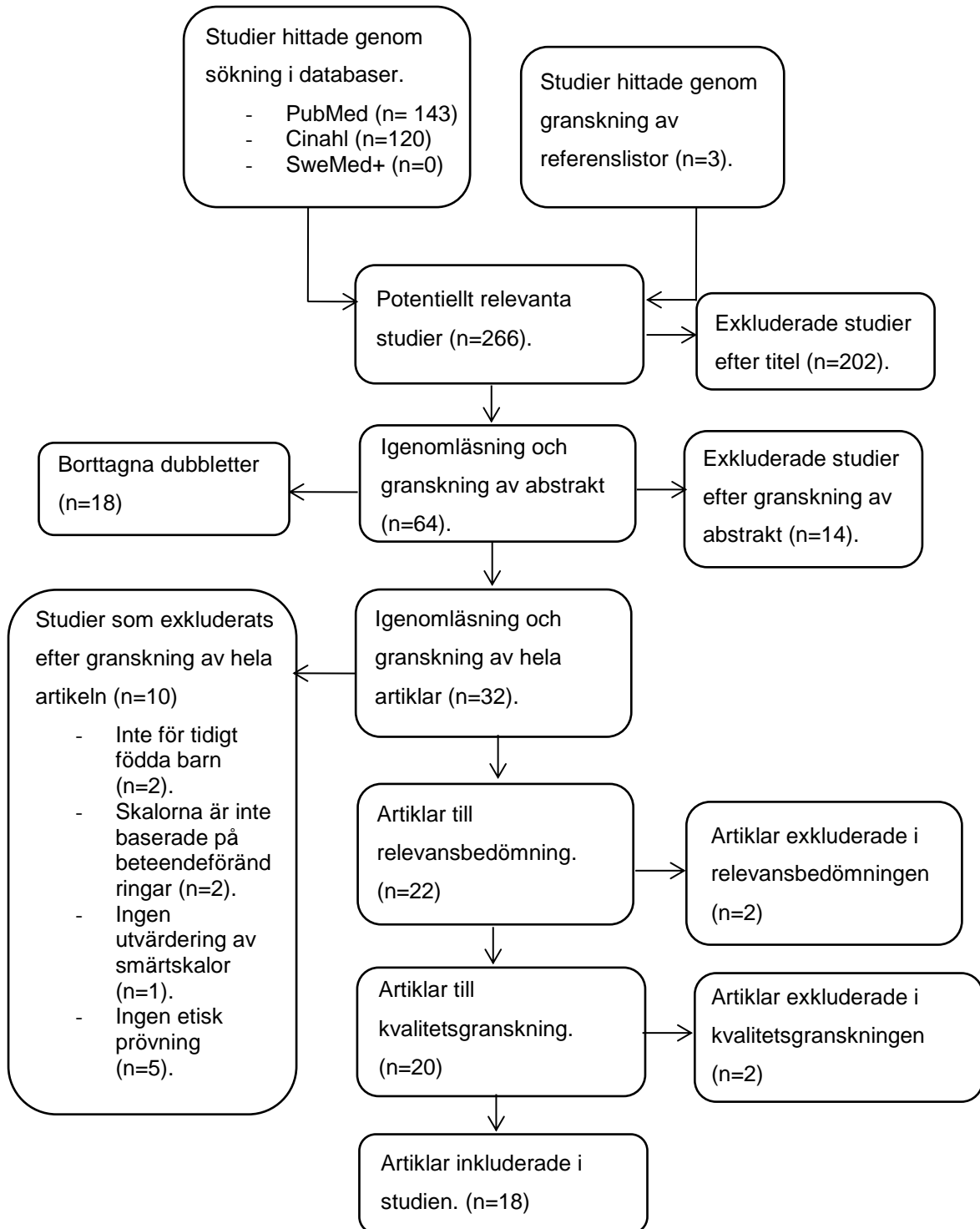
(*) = Resultatartiklar

Tabell 1. Sökschema för sökning i PubMed, Cinahl och SweMed+.

Sök	Sökord	Begränsningar	Funna titlar	Exkl.	Exkl. r.t. titel	Lästa abstrakt	Lästa i fulltext	Inkl.
Sökning i PubMed 20190301								
#1	Premature AND infant	Årtal 2000-2019. Abstrakt och fulltext tillgängliga. Eng, sve, dan och nor.	46 981	46 981	0	0	0	0
#2	Preterm AND infant	Som ovan.	34 085	34 085	0	0	0	0
#3	Neonatal AND prematurity	Som ovan.	11 107	11 107	0	0	0	0
#4	#1 AND #2 AND #3	Som ovan.	59 317	59 317	0	0	0	0
#5	Pain AND assessment	Som ovan.	63 498	63 498	0	0	0	0
#6	Pain AND measurement	Som ovan.	77 701	77 701	0	0	0	0
#7	#5 AND #6	Som ovan.	77 701	77 701	0	0	0	0
#8	#4 AND #7	Som ovan.	422	422	0	0	0	0
#9	Pain AND assessment AND scale	Som ovan.	14 924	14 924	0	0	0	0
#10	Pain AND measurement AND scale	Som ovan.	25 382	25 382	0	0	0	0
#11	#9 AND #10	Som ovan.	33 783	33 783	0	0	0	0
#12	#4 AND #11	Som ovan.	143	-	107	36	27	18
Sökning i Cinahl 20190302								
#1	Preterm infants OR preterm neonates OR preterm babies	Årtal 2000-2019. Abstrakt och fulltext tillgängliga. Eng, sve, dan och nor.	16 797	16 797	0	0	0	0
#2	Premature infants OR premature neonates OR premature babies	Som ovan.	22 770	22 770	0	0	0	0
#3	#1 AND #2	Som ovan.	27 205	27 205	0	0	0	0

#4	Pain assessment scale OR pain measurement scale	Som ovan.	24 660	24 660	0	0	0	0
#5	#3 AND #4	Som ovan.	120	-	95	25	20	14
Sökning i SweMed+ 20190302								
#1	Premature AND infant	Årtal 2000-2019. Abstrakt och fulltext tillgängliga. Eng, sve, dan och nor.	0	0	0	0	0	0
#2	Preterm AND infant	Som ovan.	28	28	0	0	0	0
#3	Neonatal AND prematurity	Som ovan.	141	141	0	0	0	0
#4	#1 AND #2 AND #3	Som ovan.	165	165	0	0	0	0
#5	Pain AND assessment AND scale	Som ovan.	16	16	0	0	0	0
#6	Pain AND measurement AND scale	Som ovan.	24	24	0	0	0	0
#7	#5 AND #6	Som ovan.	27	27	0	0	0	0
#8	#4 AND #7	Som ovan.	0	0	0	0	0	0

Figur 1. Flödesschema över litteratursökningen



Figur 1. Flödesschema över litteratursökningen.

Tabell 2. Presentation av artiklarna som ingår i litteraturstudien.

Författare, årtal, titel och ursprungsland	Skala	Design och urval	Inklusions- och exklusionskriterier	Validitet	Reliabilitet	Artikelkvalitet (poäng enligt CASP)
Caljouw et al., 2007, Measurement of pain in premature infants with a gestational age between 28 to 37 weeks: validation of the adapted COMFORT scale. Nederländerna.	COMFORT-neo	Prospektiv kohortstudie. Bekvämlighetsurval (n=57). Spädbarnen var uppdelade i fem grupper baserade på gestationsålder (v. 28–29, 30–31, 32–33, 34–35 och 36–37). 114 bedömningar.	Inkludering: a) gestationsålder mellan 28–37 veckor, b) yngre än sju dagar, c) behov att blodprovstagning i hälen och d) skriftligt tillstånd av förälder. Exkludering: barn som a) fått muskelrelaxantia eller analgetika, b) hade allvarliga neurologiska problem och c) hade en artärnål.	Kriterievaliditet: jämförelse med VAS (Pearsons rangkorrelation).	Inter-rater reliabilitet (Cohens κ). Intern konsistens (Cronbachs α).	7
Cignacco et al., 2004, Pain assessment in the neonate using the Bernese pain scale for neonates. Nederländerna.	BPSN	Prospektiv kohortstudie. Bekvämlighetsurval (n=12). Spädbarnen var uppdelade i två grupper baserade på gestationsålder (v. 27–31 och 32–41). 288 bedömningar.	Inkludering: för tidigt födda och fullgångna barn. Exkludering: barn som a) fått sedering eller analgetika vid försöket, b) blivit diagnosticerade med hjärnblödning via ultraljud, c) hade allvarliga, livshotande	Begreppsvaliditet: hur bra skalan kunde särskilja mellan olika faser i en smärtsam procedur (10 min efter matning, under värmning av foten med en varm handduk, kapillär provtagning i hälen och	Inter-rater reliabilitet (Cronbachs α). Intra-rater reliabilitet (Cronbachs α).	6

			missbildningar, d) hade saknade eller hade nedsatt sensitivitet och e) genomgått någon operation.	15 min efter; variance analysis). Kriterievaliditet: jämförelse med PIPP och VAS (Cronbachs α).		
Desai et al., 2017, Comparison of neonatal pain, agitation, and sedation scale with premature infant pain profile for the assessment of acute prolonged pain in neonates on assisted ventilation: a prospective observational study. Indien.	N-PASS	Prospektiv kohortstudie. Konsekutivt urval (n=42). Barnen var födda i gestationsvecka 28–35.	Inkludering: barn som vårdats i nonvasiv eller invasiv ventilator i >48 h. Exkludering: barn som hade a) dödliga missbildningar och b) allvarlig encefalopati.	Kriterievaliditet: jämförelse med PIPP (Spearman ρ).	-	5
Gibbins et al., 2014, Validation of the premature infant pain profile-revised (PIPP-R). Kanada.	PIPP-R	Prospektiv kohortstudie. Bekvämlighetsurval (n=202). Spädbarnen var uppdelade i tre grupper baserade på gestationsålder (v. 26–31, 32–36 och ≥ 37). 246 bedömningar av 195 olika par av bedömare (sjuksköterskor och experter).	Inkludering: barnen a) vårdades på neonatal-IVA eller neonatal vårdavdelning, b) hade ≥ 26 veckors gestationsålder vid födelsen och c) var i behov av en smärtsam och en icke-smärtsam procedur inom 24 h. Exkludering: barnen fick muskelrelaxantia.	Begreppsvaliditet: hur bra skalan kunde särskilja mellan en smärtsam (kapillär provtagning i hälen, venös provtagning, sugning av tracheostomin) och en icke-smärtsam procedur (vändning, blöjbyte, temperaturmätning;	Inter-rater reliabilitet (Pearsons korrelationskoefficient).	6

Hand et al., 2010, COVERS neonatal pain scale: development and validation. USA.	COVERS	Prospektiv kohortstudie. Konsekutivt urval (n=21). Barnen var födda i gestationsvecka 27–40.	Inkludering: alla för tidigt födda barn på avdelningen. Exkludering: barn som hade a) medfödda missbildningar, b) allvarliga neurologiska missbildningar och c) fått analgetika inom de 12 h som studien pågick.	generalised estimating equation). Kriterievaliditet: jämförelse med PIPP (Pearsons korrelationskoefficient). Begreppsvaliditet: hur bra skalan kunde särskilja mellan en smärtsam (kapillär provtagning i hälen) och en icke- smärtsam procedur (blöjbyte; Wilcoxon signed-tank test).	-	5
Holsti & Grunau, 2007, Initial validation of the behavioral indicators of pain (BIIP). Kanada	BIIP	Prospektiv kohortstudie. Bekvämlighetsurval (n=92). Barnen var födda i gestationsvecka 24–31. Totalt 100 bedömningar.	Inkludering: spädbarn a) födda i gestationsålder 24–31 veckor, och b) som vårdades på en utvald neonatal-IVA. Exkludering: spädbarn som a) fått sedering eller analgetika inom 72 h innan bedömningen, b) hade stora medfödda missbildningar och c) vars moder använt droger under graviditeten.	Kriterievaliditet: jämförelse med NIPS samt förändringar i hjärtfrekvens (Spearman rangkorrelation). Begreppsvaliditet: hur bra skalan kunde särskilja mellan tre faser (innan, under och efter) i en smärtsam procedur (kapillär provtagning i hälen; ANOVA).	Inter-rater reliabilitet (ICC). Intern konsistens (Cronbachs α).	8
Holsti et al., 2008. Is it painful or not? Discriminant validity	BIIP	Prospektiv kohortstudie. Bekvämlighetsurval (n=69).	Inkludering: spädbarn a) födda i gestationsålder 24–32	Kriterievaliditet: jämförelse förändringar i hjärtfrekvens (ANOVA).	-	8

of the behavioural indicators of infant pain (BIIP) scale. Kanada.		Barnen var födda i gestationsvecka 24–32. Totalt 100 bedömningar.	veckor, och b) som vårdades på en utvald neonatal-IVA. Exkludering: spädbarn som a) fått sedering eller analgetika inom 72 h innan bedömningen, b) hade stora medfödda missbildningar och c) vars moder använt droger under graviditeten.	Begreppsvaliditet: hur bra skalan kunde särskilja mellan tre faser (innan, under och efter) i en smärtsam procedur (kapillär provtagning i hälen) samt mellan en smärtsam (kapillär provtagning i hälen) och en icke-smärtsam (blöjbyte) procedur (ANOVA).		
Huang et al. (2018), Evaluation of three pain assessment scales used for ventilated neonates. Kina.	PIPP-R N-PASS NIAPAS	Prospektiv kohortstudie. Bekvämlighetsurval (n=90). Spädbarnen var uppdelade i tre grupper baserade på gestationsålder (<v. 28, v. 28–36, och v. 37 och äldre). 1080 bedömningar av 40 olika bedömare.	Inkludering: för tidigt födda barn på en avdelning som a) var intuberade och vårdades i ventilator, b) var 1–30 dagar gamla, och c) behövde byta blöjor och få blodprovstagning. Exkludering: blivit diagnostiserade med a) allvarlig asfyxi, b) hypoxisk-ischemisk encefalopati, c) förlamning, d) allvarliga medfödda missbildningar, e) nära döden, och f) vårdades i CPAP.	Kriterievaliditet: jämförelse mellan de tre skalorna (Spearman rangkorrelation). Begreppsvaliditet: jämförelse mellan icke-smärtsam (blöjbyte) och smärtsam procedur (kapillär provtagning i hälen; Friedmans test).	Inter-rater (ICC). Intern konsistens (Cronbachs α).	6
Hummel et al., 2008,	N-PASS	Prospektiv kohortstudie. Bekvämlighetsurval (n=46).	Inkludering: intuberade och/eller postoperativa för	Kriterievaliditet: jämförelse med PIPP	Inter-rater reliabilitet (ICC).	5

Clinical reliability and validity of the N-PASS: neonatal pain, agitation and sedation scale with prolonged pain. USA.		Barnen var födda i gestationsvecka 23–40. 72 bedömningar av 10 olika bedömare.	tidigt födda barn som fick farmakologisk behandling (lugnande, opioider eller båda). Exkludering: -	(Spearman rangkorrelation). Begreppsvaliditet: jämförelse mellan innan och efter farmakologisk behandling (Wilcoxon test).	Intern konsistens (Cronbachs α).	
Hummel et al. (2010) Validity and reliability of the N-PASS assessment tool with acute pain. USA.	N-PASS	Prospektiv kohortstudie. Bekvämlighetsurval (n=42). Barnen var födda i gestationsvecka 23–42. 59 bedömningar av olika par med bedömare (sjuksköterska och expert).	Inkludering: för tidigt födda barn på en avdelning som a) var födda i gestationsvecka 23–42 och b) var >30 dagar gamla. Exkludering: a) kritiskt eller instabilt tillstånd där proceduren bedömdes öka stressen, b) farmakologisk förlamning, c) för hög sedering och d) neurologiska tillstånd som ger onormal smärtrespons, t. ex spina bifida och hypoxisk-ischemisk encefalopati.	Kriterievaliditet: för att se hur väl de olika skalorna stämmer överens (Spearman ρ korrelation). Begreppsvaliditet: jämförelse mellan en äkta kapillär provtagning i hälen och en falsk (Wilcoxon test).	Inter-rater reliabilitet (ICC). Intra-rater (Pearsons korrelationskoefficient). Intern konsistens (Cronbachs α).	5
Kappesser et al. (2019). Pain-specific reactions or indicators of a general stress response? Investigating the discriminant validity of 5 well-	BPSN NFCS-R NIPS N-PASS PIPP-R	Prospektiv kohortstudie. Bekvämlighetsurval (n=42). Barnen var födda i gestationsvecka <37. Två olika bedömare gjorde bedömningarna.	Inkludering: för tidigt födda vårdade på en neonatal-IVA. Exkludering: barn som hade åkommor som ansågs kunna påverka de indikatorer som ingick i smärtskattnings-skalorna, t.ex frakturer,	Kriterievaliditet: jämförelse mellan alla fem skalorna (Pearson korrelationskoefficient). Begreppsvaliditet: jämförelse mellan en icke-smärtsam (blöjbyte)	Inter-rater reliabilitet (ICC). Intern konsistens (Cronbachs α).	7

established neonatal pain assessment tools. Germany.			neurologiska sjukdomar, hjärnblödning, hepatorenalt syndrom, ansiktsmissbildningar, förändrad tonus, samt mödrar med alkohol- eller narkotikamissbruk.	och en smärtsam (venös blodprovstagning) procedur (t-test).		
Liaw et al. (2011), Psychometric analysis of a Taiwan-version pain assessment scale for preterm infants. Taiwan.	PASPI	Prospektiv kohortstudie. Bekvämlighetsurval (n=60). Barnen var födda i gestationsvecka 27–37. Två olika bedömare gjorde 240 bedömningarna.	Inkludering: nyfödda barn a) födda i gestationsvecka 27–37, b) ålder <30 dagar och c) ≤12 poäng på Neonatal therapeutics intervention scoring system (NTISS). Exkludering: om spädbarnen a) hade farmakologisk behandling med lugnande, analgetika eller muskelrelaxantia, b) hade >grad II hjärnblödningar, c) var intuberade och vårdades i ventilator och d) hade continuous positive airway pressure (CPAP).	Innehållsvaliditet: indikatorerna som ingick i PASPI utvärderades av tio neonatala, kliniska experter (fem neonatologer och fem sjuksköterskor; CVI). Kriterievaliditet: jämförelse med PIPP och VAS (ICC). Begreppsvaliditet: hur bra skalan kunde särskilja mellan fyra faser (innan, under, tre och tio minuter efter) i en smärtsam procedur (kapillär provtagning i hälen; ANOVA).	Inter-rater (ICC). Intern konsistens (Cronbachs α).	6
Lundqvist et al. (2014), Development and psychometric properties of the	ALPS-neo	Klinisk observationsstudie. Bekvämlighetsurval (n=40). Nyfödda barn födda i gestationsvecka 25–42.	-	Innehållsvaliditet: indikatorerna som ingick i ALPS-neo utvärderades både internt (25	Inter-rater (ICC). Intern konsistens (Cronbachs α)	6

Swedich ALPS-neo pain and stress assessment scale for newborn infants. Sweden.		I ett första steg gjordes 125 bedömningar av 10 par med sjuksköterskor på en svensk neonatal-IVA och i ett andra steg 96 bedömningar av 12 par sjuksköterskor på en svensk neonatal vårdavdelning.		sjuksköterskor och en neonatolog) och externt (tre sjuksköterskor som doktorerat i omvårdnad) i ett första steg, och av fem sjuksköterskor i ett andra steg där man studerade videosekvenser med spädbarn och noterade varför de bedömde som de gjorde.		
McNair et al. (2004), Postoperative pain assessment in the neonatal intensive care unit. Kanada.	CRIES PIPP	Prospektiv kohortstudie. Bekvämlighetsurval (n=51). Tre olika personer (två sjuksköterskor och en expert) gjorde bedömningar oberoende av varandra.	Inkludering: nyfödda barn a) med en gestationsålder mellan 28 och 42 veckor, b) yngre än 30 dagar, c) genomgått en operation och d) inte ha några neurologiska missbildningar eller sjukdomar.	Kriterievaliditet: för att se hur väl de olika skalorna stämmer överens (covariance matrix metod).	-	7
O'Sullivan et al. (2016), The validity and clinical utility of the COVERS scale and pain assessment tool for assessing pain in neonates admitted to an intensive care unit. Australien.	COVERS PAT	Prospektiv kohortstudie. Bekvämlighetsurval (n=80). 72 sjuksköterskor gjorde bedömningarna med båda smärtskattningsskalorna.	Inkludering: spädbarn i alla gestationsåldrar. Exkludering: spädbarn som a) var i kritiskt tillstånd eller förväntades dö, b) hade neurologiska skador/sjukdomar, c) fick muskelrelaxantia, d) var farmakologiskt förlamade och e) var så sederade att deras	Kriterievaliditet: jämförelse mellan de två skalorna och moderns smärtbedömning (0–10 poäng; Spearman korrelation). Begreppsvaliditet: jämförelse mellan ett icke-smärtsam (blöjbyte) och en smärtsam (kapillär provtagning i hälen;	Inter-rater reliabilitet (κ measurement of agreement och ICC). Intern konsistens (Cronbach's α).	7

			smärtresponser var reducerade.	Wilcoxon signed ranks tests).		
Pölkki et al. (2001), Development and preliminary validation of the Neonatal infant acute pain assessment scale (NIAPAS). Finland.	NIAPAS	Prospektiv kohortstudie. Konsekutivt urval (n=34). Spädbarn födda mellan gestationsålder 23 och 42 veckors), 180 bedömningar av två olika bedömare, som även gjorde om samma bedömning 2–3 veckor senare.	Inkludering: Alla för tidigt födda barn på avdelningen. Exkludering: Barn som a) fått sederung eller analgetika vid tidpunkten av smärtskattning, b) hade tecken på hjärnskada, och c) som var nära döden.	Innehållsvaliditet: en panel med experter utvärderade de olika bedömningspunkterna i skalan. Kriterievaliditet: jämförelse med NIPS (Pearson korrelationskoefficient). Begreppsvaliditet: hur bra skalan kunde särskilja mellan olika faser (innan, under och efter) i en smärtsam procedur (blodprov eller sugning i trachen; ANOVA och post hoc test med Bonferroni).	Inter-rater reliabilitet: (Pearson korrelationskoefficient). Intra-rater reliabilitet (Pearson korrelationskoefficient). Intern konsistens (Cronbachs α).	6
Spence et al. (2005), A reliable pain assessment tool for clinical assessment in the neonatal intensive care unit. Australien.	PAT	Prospektiv kohortstudie. Konsekutivt urval (n=144). Tolv sjuksköterskor gjorde bedömningarna.	Inkludering: alla spädbarn (både för tidigt födda och fullgångna) som vårdades på två neo-IVA. Exkludering: spädbarn som fått muskelrelaxantia.	Kriterievaliditet: jämförelse med CRIES och med moderns bedömning av barnets smärta (VAS; Pearson korrelationskoefficient).	Inter-rater reliabilitet (ICC).	7
Woragidpoonpol et al. (2018), Development	CPSPN	Prospektiv kohortstudie. Bekvämlighetsurval (n=19).	Inkludering: spädbarn som a) vårdades på neonatal-IVA, b)	Kriterievaliditet: jämförelse med PIPP-R	Inter-rater reliabilitet (ICC).	7

of a clinical pain scale for preterm neonates. Thailand.	53 bedömningar av två varandra oberoende bedömare.	var födda i gestationsvecka 22–36, c) var planlagda att genomgå en smärtsam procedur inom 24 h och d) föräldrarnas skriftliga tillstånd. Exkludering: -	(Spearman's ρ korrelation). Begreppsvaliditet: hur bra skalan kunde särskilja mellan olika faser (innan, under och efter) i en smärtsam procedur (kapillärprovtagning i hälen; Kruskal-Wallis test).	Intern konsistens (Cronbach's α). Item analysis: beräknar korrelationen mellan olika indikatorer i skalan.
--	--	--	--	---

Tabell 3. Översikt över de fjorton smärtskattningsskalor som identifierats i studien.

Smärtskattningsinstrument	Anpassad för procedur- eller pågående smärta	Antal bedömningspunkter, poängskala.	Fysiologisk indikator	Beteendeindikator
Astrid Lindgren children's hospital pain scale-neo (ALPS-neo)	Pågående	Fem indikatorer. Alla ger 0–2 poäng. 0–10 poäng.	-	Ansiktsuttryck, andnings-mönster, extremitetstonus, hand- och fotrörelser och aktivitet.
Behavioral indicators of infant pain scale (BIIP)	Procedur	Åtta indikatorer. Vakenhetsgrad ger 0–2 poäng, övriga 0–1. 0–9 poäng.	Hjärtfrekvens och SaO ₂ .	Vakenhetsgrad, fem ansiktsuttryck (rynkande ögonbryn, ihoppressade ögon, näsläpp-fåra, stäng mun och synlig tunga) och två handrörelser.
Bernese pain scale neonates (BPSN)	Procedur	Nio indikatorer. Alla indikatorer ger 0–3 poäng. 0–27 poäng.	Hjärtfrekvens, andningsfrekvens och SaO ₂ .	Sömn/vakenhet, gråt, hur lång det tar att bli lugn, hudfärg, ansiktsuttryck och hållning.
Clinical pain scale for preterm neonates (CPSPN)	Procedur	Fyra indikatorer. Man delar upp de för tidigt födda barnen i två grupper baserat på gestationsålder. Alla indikatorer ger 0–2 poäng. 0–8 poäng.	Ökning av hjärtfrekvens.	Övre ansiktsuttryck, nedre ansiktsuttryck och sömn/vakenhet.
COMFORTneo	Pågående	Åtta indikatorer. Alla ger 0–5 poäng (0 poäng vid inget lidande och 5 poäng vid mycket lidande) 0–40 poäng.	Hjärtfrekvens.	Vakenhet, upprördhet, andningsmönster, gråt, generella rörelser, muskeltonus, och ansiktsuttryck.
COVERS scale	Procedur	Sex indikatorer. Alla ger 0–2 poäng. 0–10 poäng.	Syrgasbehov, ökad hjärtfrekvens och puls.	Ansiktsuttryck, gråt, oro och sömnlöshet.

Crying, requires oxygen, increased vital signs, expression and sleepless measure (CRIES)	Pågående	Fem indikatorer. Alla ger 0–2 poäng. 0–12 poäng.	Syrgasbehov, ökad hjärtfrekvens och puls.	Ansiktsuttryck, gråt och sömnlöshet.
Neonatal infant acute pain assessment scale (NIAPAS)	Procedur	Åtta indikatorer, som ger en två, tre eller fyra poäng. 0–18 poäng.	Andning, hjärtfrekvens och SaO ₂ .	Vakenhet, ansiktsuttryck, gråt, muskeltonus, och reaktion vid hantering.
Neonatal infant pain scale (NIPS)	Procedur	Sex indikatorer. Alla ger 0 (avslappnad) eller 1 (beteendet är närvarande) förutom gråt som ger 0–2 poäng (ingen gråt, gnällande, ordentlig gråt). 0–7 poäng.	Andningsmönster	Ansiktsuttryck, gråt, armrörelser, benrörelser, vakenhet
Neonatal facial activity coding scale (NFCS-R)	Procedur	Fem indikatorer. Alla ger 0 (beteendet finns inte) eller 1 poäng (beteendet finns), eller så skriver man nv (not visible). Ger 0–5 poäng.	-	Rynkande ögonbryn, ihoppressade ögon, näsläpp-fåra, stäng mun och synlig tunga.
Neonatal Pain, agitation and sedation scale (N-PASS)	Procedur och pågående smärta	Fem indikatorer. Poäng från -2 till +2 där 0 innebär ingen smärta eller sedering. Positiva värden innebär smärta medan negativa värden innebär sedering. Extra poäng ges om barnet är född för tidigt, maximalt tre poäng. 0–10 poäng.	Förändring i en vital parameter, kan vara hjärtfrekvens, blodtryck, puls och SaO ₂ .	Ansiktsuttryck, gråt, extremiteters tonus, aktivitet/vakenhet/sömn, och hand- och fotrörelser.
Pain assessment scale for preterm infants (PASPI)	Procedur			

Pain assessment tool (PAT)	Procedur	Tio indikatorer. Förutom fysiologiska och beteendeindikatorer tas hänsyn till sjuksköterskans uppfattning. Alla ger 0–2 poäng. 0–20 poäng.	Andningsfrekvens, hjärtfrekvens, blodtryck och SaO ₂ .	Kroppshållning, gråt, sovmonster, ansiktsuttryck, färg.
Premature infant pain profile-R (PIPP-R)	Procedur och pågående.	Två fysiologiska och tre beteendeindikatorer, samt även gestationsålder och vakenhetsstadium (aktivt/tyst/vaken/sovande). 0–21 poäng.	Hjärtfrekvens och SaO ₂ .	Rynkande ögonbryn, ihoppressade ögon, och näsläppfåra.