

# Hoitotyön interventioiden vaikutusten arviointi efektikokojen ja meta-analyysin avulla

**ARI VOUTILAINEN**

FT, dosentti, projektitutkija

Itä-Suomen yliopisto  
Hoitotieteen laitos

**MARJA HÄRKÄNEN**

TtT, tutkijatohtori

Itä-Suomen yliopisto  
Hoitotieteen laitos

## TIIVISTELMÄ

Intervention vaikutusta tulisi arvioida efektikokojen ja meta-analyysin avulla. Tämän tutkimuksen tarkoituksena on kuvata esimerkkejä efektikokojen laskemisesta, efektikokojen käyttämisestä meta-analyysissä osana hoitotyön interventioiden vaikutusten arviointia ja meta-analyysin tulosten tulkintaa. Esimerkkiaineistona käytetään aikaisemmin systemaattista kirjallisuuskatsausta varten kerättyä aineistoa. Aineistoon kuuluvissa tutkimuksissa (n=14) oli testattu sairaanhoitajien lääkehoidon osaamista ja turvallisuutta parantamaan pyrkiviä koulutuksellisia interventiota. Yhdessäkään alkuperäistutkimuksessa intervention vaikutusta ei ollut arvioitu efektikokojen avulla, vaan johtopäätökset oli tehty tilastollisen merkitsevyyden perusteella. Aineistossa intervention vaikutus oli yhteydessä interventiota toteutuksen ja sen arvioinnin väliseen aika-vaiveeseen. Havainto tähdentää intervention välittömän vaikutuksen ja sen varsinaisen lopputuloksen, vaikuttavuuden, erottamista toisistaan. Meta-analyysin tulosta on arvioitava sekä yhdistettävien tutkimusten toiminnallisen johdonmukaisuuden että tunnusluvuilla, kuten Cochran's Q, I<sup>2</sup> ja T<sup>2</sup>, ilmaistavan vastemuuttujan vaihteluun perustuvan numeerisen heterogeenisuuden perusteella. Esitellyt Cohenin ja Hedgesin efektikoot ovat käyttökelpoisia erityisesti, kun tarkoituksena on muuntaa erilaisia vaikutuksen tunnuslukuja "yleispäteväksi" efektikooksi ja siten mahdollistaa tutkimusten yhdistäminen meta-analyttisesti. Tätä artikkelia voi käyttää käsikirjana efektikokoja laskettaessa ja meta-analyysien tuloksia tulkittaessa.

Avainsanat: interventio, meta-analyysi, tilastollinen analyysi

## ABSTRACT

### **Evaluation of effects of nursing interventions by means of effect sizes and meta-analysis**

*Ari Voutilainen, PhD, Adjunct professor, Project researcher*

*Marja Härkänen, PhD, Postdoctoral researcher*

The effect of an intervention should be evaluated by means of effect sizes and meta-analysis. The purpose of this study is to describe examples of effect size calculations and their use in a meta-analysis as a part of the evaluation of the effects of nursing interventions as well as interpretation of results of a meta-analysis. A dataset previously collected for a systematic literature review is used as an example dataset. Educational interventions performed to improve medication administration skills of registered nurses and increase the safety of medication administration were tested in studies (n=14) belonging to the dataset. The effects of the interventions were not evaluated by means of effect sizes in any of the original studies, but conclusions were drawn on the basis of statistical significance. In the dataset, the effect of an intervention was related to the time lag between the execution of the intervention and its evaluation. The finding emphasizes distinguishing the immediate effect of an intervention from its actual outcome, effectiveness. A result of a meta-analysis needs to be evaluated based on both the functional consistency of the studies combined and numerical heterogeneity expressed by statistics, such as Cochran's Q, I<sup>2</sup>, and T<sup>2</sup>. The presented Cohen's and Hedges' effect sizes are useful; especially, when the purpose is to transform different effect statistics to a "general" effect size and thus enable a meta-analytic combination of studies. This article can be used as a manual when calculating effects sizes and interpreting results of meta-analyses.

Key words: intervention study, meta-analysis, statistical data analysis

Saapunut 15.05.2016

Hyväksytty julkaistavaksi 21.12.2016

**Mitä tutkimusaiheesta jo tiedetään?**

- Intervention vaikutusta tulisi arvioida efektikokojen ja meta-analyysin avulla.
- Hoitotieteessä on tehty vähän meta-analyysejä.

**Mitä uutta tietoa artikkeli tuo?**

- Tässä artikkelissa esitetään useita efektikokojen ja meta-analyysin arviointiin liittyvien tunnuslukujen laskentakaavoja ja jokaisen kaavan soveltaminen havainnollistetaan esimerkin avulla.
- Tätä artikkelia voi käyttää käsikirjana efektikokoja laskettaessa ja meta-analyysien tuloksia tulkittaessa.

**Mikä merkitys tutkimuksella on hoitotyölle, hoitotyön koulutukselle ja johtamiselle?**

- Interventioiden implementoinnissa voidaan keskittyä toimivimpiin ja tehokkaimpiin interventioihin, kun niiden vaikutusta arvioidaan efektikokojen ja meta-analyysien avulla.
- Meta-analyysit vahvistavat hoitotieteen tutkimuksen laatua tarjoamalla korkeinta mahdollista näyttöä interventioiden vaikutuksesta.

**Tutkimuksen lähtökohdat**

Tutkimuksen laatua arvioidaan tutkimusraportin perusteella laadunarviointityökaluilla, joita ovat kehittäneet esimerkiksi Joanna Briggs Instituutti (JBI) (HOTUS 2016), The Cochrane Collaboration (Higgins & Green 2011) ja Effective Public Health Practice Project (EPHPP 2016). Vaikka laadunarviointityökaluilla saadaan kokonaisarvio arvioitavan tutkimuksen laadusta (Zeng ym. 2015), sen perusteella ei voi tehdä johtopäätöksiä tutkimuksessa suoritetun intervention tilastollisesta merkitsevyydestä, vaikutuksesta tai vaikuttavuudesta. Intervention tilastollinen merkitsevyys selvitetään tilastollisten testien ja niiden testisuureille laskettujen p-arvojen avulla. Intervention vaikutuksen suuruutta puolestaan arvioidaan laskemalla efektikokoja ja tekemällä meta-analyysejä (Sullivan & Feinn 2012). Meta-analyysillä tarkoitetaan yksittäisten toisistaan riippumattomien määrällisten tutkimusten numeerista yhdistämistä eli yhdistetyn efektikoon laskemista tilastollisin menetelmin (Chamberlain 2007). Meta-analyysi on siis eri asia kuin metasynteesi, jolla tarkoitetaan laadullisten tutkimusten tulosten synteesiä (Korhonen ym. 2013).

Käsitteinä intervention vaikutus (engl. effect) ja vaikuttavuus (engl. effectivity tai ef-

fectiveness) eivät ole synonyymejä, vaan ne tyypillisesti viittaavat erilaisiin tulosmuuttujiin. Esimerkiksi lääke voi olla vaikuttava, vaikka sillä ei olisi vaikuttavuutta. Toivotulla tavalla muuttuvat potilaiden veriarvot, joilla mitataan lääkkeen vaikutusta, eivät automaattisesti tarkoita pienentynyttä kuolleisuutta, jolla mitataan lääkkeen vaikuttavuutta. Vaikuttavuus siis liitetään intervention varsinaiseen lopputulemaan (engl. outcome), ja useimmiten sen arviointi vaatii pidempää seurantaa kuin vaikutuksen arviointi.

Meta-analyysissä yksittäisissä tutkimuksissa suoritetuille interventioille lasketut efektikoot yhdistetään painotetuksi keskiarvoksi. Yksittäisen intervention efektikoon painokerrotimeen vaikuttavat alkuperäisen aineiston koko ja tutkittavan muuttujan vaihtelu. Vaihtelulla tarkoitetaan käytettävästä meta-analyysimallista riippuen joko pelkästään tutkittavan vastemuuttujan vaihtelua tutkimusten sisällä tai sekä muuttujan vaihtelua tutkimusten sisällä että sen vaihtelua yhdistettävien tutkimusten välillä. Vaihtelun mittana käytetään varianssia ja tutkimuksen painokertoimenä meta-analyysissä käänteisvarienssia. (ks. esim. Borenstein ym. 2009.) Myös tilastollisesti merkityksetön tai hypoteesin vastainen tulos on meta-analyysin näkökulmasta tarkasteltavan ilmiön ymmärtämistä lisäävä (Schmidt 1992, Borenstein ym. 2009).

Perinteisesti meta-analyysiä ja systemaattista kirjallisuuskatsausta on pidetty vaihtoehtoisina menetelminä näytön asteen määrittämisessä (Fineout-Overholt & Johnston 2005), mikä ei johda parhaaseen mahdolliseen käytäntöön, sillä laadullinen systemaattinen kirjallisuuskatsaus ja määrällinen meta-analyysi eivät tuota toisiaan korvaavaa vaan täydentävää tietoa (Chamberlain 2007). Vahvin näyttö saadaan yhdistämällä systemaattinen kirjallisuuskatsaus ja meta-analyysi, ts. ne ovat vahvinta tutkimusnäyttöä tuottavan tutkimusprosessin vaiheita, systemaattinen kirjallisuuskatsaus edeltää meta-analyysiä (Shorten & Shorten 2013, Malmivaara & Komulainen 2014).

Hoitotieteessä on tehty vähän meta-analyysyjä. CINAHL-viitetietokannasta löytyy 271 vuosina 1982–2016 hoitotieteellisessä lehdessä (lehden nimessä sana: nurs\*) julkaistua tieteellistä vertaisarvioitua artikkelia, jonka otsikossa mainitaan meta-analyysi (haku tehty 29.3.2016). Ero esimerkiksi lääketieteen on suuri. Vastaavilla hakuehdoilla CINAHL-viitetietokannasta löytyy noin 8500 lääketieteellisessä lehdessä (lehden nimessä sana: medic\*) julkaistua meta-analyysiartikkelia. Tämä menetelmäartikkeli on tarpeellinen, koska se osaltaan edesauttaa meta-analyysin käyttöä hoitotieteessä kuvaamalla esimerkin avulla meta-analyysin tekemisen eri vaiheita.

## Tutkimuksen tarkoitus

Tämän artikkelin tarkoituksena on yksityiskohtaisesti kuvata esimerkkejä erilaisten efektikokojen laskemisesta, efektikokojen käyttämisestä meta-analyysissä osana hoitotyön interventioiden vaikutusten arviointia ja meta-analyysin tulosten tulkintaa. Artikkeliiin on koottu efektikokojen ja meta-analyysin arviointiin liittyvien tunnuslukujen laskenta-kaavoja useista eri lähteistä ja jokaisen kaavan soveltaminen on havainnollistettu konkreettisen esimerkin avulla. Artikkelin tavoit-

teena on lisätä efektikokojen ja meta-analyysin käyttöä hoitotieteellisessä tutkimuksessa.

## Aineisto

Tässä tutkimuksessa käytetään samaa aineistoa kuin vuonna 2016 Nurse Education Today -lehdessä julkaistussa systemaattisessa kirjallisuuskatsauksessa ja meta-analyysissä (Härkänen ym. 2016). Koska sairaanhoitajien toteuttamassa lääkehoidossa on havaittu tapahtuvan paljon vaaratapahtumia, alkuperäisen tutkimuksen tarkoituksena oli kartoittaa sairaanhoitajien lääkehoidon osaamista ja lääkehoidon turvallisuutta kehittäviä koulutusinterventioita sekä arvioida niiden laatua ja vaikuttavuutta. CINAHL-, PubMed-, Scopus-, Cochrane-, PsycInfo- ja Medic-viitetietokannoista kerättiin systemaattisesti aineisto, joka koostui tieteellisissä vertaisarvioituissa englanninkielisissä lehdissä vuosina 2006–2014 julkaistusta tutkimuksista. Alkuperäisaineistoon hyväksyttiin sisäänotto-kriteereiden mukaisesti 1) vertaisarvioidut suomeksi tai englanniksi julkaistut interventiotutkimukset, 2) tutkimukset, jotka koskivat sairaanhoitajan toteuttamana minkä tahansa lääkkeen antamista potilaille sekä 3) tutkimukset, jotka oli toteutettu sairaaloissa mukaan lukien kaikki osastot ja erikoisalat. Aineistosta hylättiin poissulkukriteereiden mukaisesti 1) tutkimukset, joissa tutkittiin muita kuin koulutusinterventioita, 2) tutkimukset, joissa lääkkeen antaminen toteutettiin muun henkilön kuin sairaanhoitajan toimesta, 3) sairaaloiden ulkopuolella toteutetut tutkimukset sekä 4) harmaa kirjallisuus mukaan lukien opinnäytetyöt ja konferenssijulkaisut. Haussa löydetyistä 755 tutkimuksesta 726 hylättiin otsikon ja/tai tiivistelmän perusteella. Artikkelin kokotekstin perusteella hylättiin vielä neljä tutkimusta ja 14 duplikaattia sekä lisättiin kolme tutkimusta, jotka löydettiin läpikäymällä valittujen tutkimusten lähdeluettelot. Siten lopullinen aineisto koostui 14:sta tutkimuksesta.

Yleisesti, tieteenalasta riippumatta, suositellaan, että meta-analyysiin valitaan ainoastaan menetelmällisesti laadukkaita tutkimuksia (Malmivaara & Komulainen 2014). Tässä artikkelissa analysoitiin kaikki systemaattisella tiedonhaulla löydetty tutkimukset (n=14), vaikka ainoastaan viisi niistä oli aiemmin EPHPP:n laadunarviointityökalun (EPHPP 2016) avulla arvioitu laadultaan kohtalaiseksi tai hyväksi (Härkänen ym. 2016). Usean tutkimuksen yhdistäminen meta-analyttisesti palveli tämän artikkelin tarkoitusta, koska se mahdollisti useampien erilaisten efektikokojen laskemisen ja monipuolisemman meta-analyysin tulosten tulkinna.

Lasketeknisistä syistä aineisto jaettiin kahteen ryhmään. Jakoa hyödynnettiin erilaisten efektikokojen laskennassa ja molemmille ryhmille tehtiin oma meta-analyysi. En-

simmäisen ryhmän (Taulukko 1) tutkimuksissa intervention onnistumista oli selvitetty suorittamalla koulutettaville sairaanhoitajille taito- tai tietotesti ennen ja jälkeen intervention ja tulokset oli ilmoitettu testipistemäärien keskiarvona. Toisen ryhmän (Taulukko 2) yhdessä tutkimuksessa (Sherriff ym. 2012) intervention onnistumista oli selvitetty taitotestin avulla, mutta tuloksissa ei ilmoitettu testipistemääriä, vaan testin läpäisseiden sairaanhoitajien prosentuaalinen osuus testiin osallistuneista. Muissa toisen ryhmän tutkimuksissa intervention onnistumista oli selvitetty havainnoimalla sairaanhoitajien toteuttamaa lääkitystä käytännön potilastilanteissa ennen ja jälkeen intervention ja tulokset oli ilmoitettu virheellisten lääkitystapahtumien prosentuaalisena osuutena kaikista lääkitystapahtumista.

Taulukko 1. Meta-analyysi tutkimuksista, joissa intervention vaikutusta oli selvitetty taito- tai tietotestillä.

Tutkimus	Testin pistemäärä		Efektikoko Cohen's d	Hedges' g <sup>b</sup>	Metsikkökuvio	Painoarvo
	Ennen <sup>a</sup>	Jälkeen <sup>a</sup>				
Simonsen ym. (2014)	11,1±1,7, 92	11,6±2,0, 68	0,27	0,27, -0,07–0,61		19,9%
Tsai ym. (2008)	13,7±2,0, 37	15,9±1,6, 37	1,23	1,19, 0,60–1,78		17,2%
Abbasinazari ym. (2012)	4,5±1,2, 200	6,2±1,2, 200	1,33	1,32, 1,06–1,59		20,5%
Lu ym. (2013)	77,2±15,5, 120	94,7±7,6, 113	1,43	1,42, 1,06–1,79		19,6%
Harne-Britner ym. (2006)	15,5±2,5, 22	18,6±1,3, 22	1,61	1,51, 0,66–2,36		14,1%
Sung ym. (2008)	57,7±5,5, 24	82,2±8,8, 24	3,42	3,24, 1,83–4,65		8,7%
<b>Yhdistetty vaikutus</b>				<b>1,30, 0,77–1,84</b>		<b>100,0%</b>

Satunnaismalli, yhdistetyn vaikutuksen testi:  $z=4,74$ ,  $p<0,0005$   
Heterogeenisuus:  $Tau^2=2,10$ ,  $Q=7,32$ ,  $vapausasteet=5$ ,  $I^2=31,7\%$

<sup>a</sup>keskiarvo±keskihajonta, n  
<sup>b</sup>efektikoko ja sen 95%:n luottamusväli

Taulukko 2. Meta-analyysi tutkimuksista, joissa intervention vaikutusta oli selvitetty havainnoimalla lääkitystapahtumia ja tuloksissa raportoitiin virheiden prosentuaalinen osuus.

Tutkimus	Lääkitysvirheitä %		Efektikoko Cohen's h	Hedges' g <sup>b</sup>	Metsikkökuvio	Painoarvo
	Ennen <sup>a</sup>	Jälkeen <sup>a</sup>				
Blank ym. (2011)	43,9, 38,2–49,7, 287	34,5, 29,2–40,0, 296	0,19	0,22, 0,04–0,40		17,2%
Otero ym. (2008)	11,4, 10,0–12,9, 1764	7,3, 6,4–8,3, 2732	0,14	0,27, 0,16–0,39		19,9%
Schneider ym. (2006)	77,4, 72,5–81,9, 301	65,6, 60,0–71,0, 285	0,26	0,32, 0,12–0,52		16,5%
Xu ym. (2014)	6,0, 4,1–8,3, 496	3,2, 1,8–4,9, 496	0,14	0,36, 0,02–0,70		11,2%
Sherriff ym. (2012)	73,7 <sup>c</sup> , 63,6–82,7, 81	56,9 <sup>c</sup> , 44,1–69,3, 58	0,36	0,42, 0,03–0,82		9,6%
Hohenhaus ym. (2008)	52,7, 33,4–71,5, 25	30,8, 10,3–56,5, 14	0,45	0,54, -0,22–1,29		3,8%
Kliger ym. (2009)	14,6, 11,9–17,5, 604	4,2, 2,8–5,9, 614	0,37	0,74, 0,49–0,99		14,5%
Ford ym. (2010)	30,8, 23,8–38,3, 156	6,2, 2,3–11,8, 97	0,67	1,05, 0,56–1,54		7,3%
<b>Yhdistetty vaikutus</b>				<b>0,43, 0,27–0,59</b>		<b>100,0%</b>

Satunnaismalli, yhdistetyn vaikutuksen testi:  $z=5,21$ ,  $p<0,0005$   
Heterogeenisuus:  $Tau^2=0,25$ ,  $Q=8,43$ ,  $vapausasteet=7$ ,  $I^2=17,0\%$

<sup>a</sup>prosenttiosuus ja sen 95%:n luottamusväli, n  
<sup>b</sup>efektikoko ja sen 95%:n luottamusväli

<sup>c</sup>taitotestin reputaneiden prosentuaalinen osuus

## Menetelmät

### *Efektikoon laskeminen ja tulkinta*

Efektikoko voidaan laskea korrelaation, ryhmien välisen eron tai jonkin suhdeluvun perusteella. Erilaiset efektikoot ovat suhteellisen helposti muunnettavissa toisiaan vastaaviksi (Rosnow & Rosenthal 2003, Borenstein ym. 2009, Lakens 2013). Tässä artikkelissa intervention vaikutukselle laskettiin Cohen's d ja Hedges' g efektikoot, jos alkuperäisessä tutkimuksessa intervention vaikutusta oli selvitetty tekemällä taito- tai tietotestejä ja niiden tulokset oli ilmoitettu pistemäärinä (Taulukko 1). Jos intervention tulokset oli ilmoitettu ryhmäkohtaisina prosentiosuuksina, laskettiin Cohen's h ja Hedges' g efektikoot (Taulukko 2). Koska Cohen's h on arkussini-muunnettujen populaatio-osuuksien erotus (Rosnow & Rosenthal 2003), laskettiin suhteellisten osuuksien arkussini-muunnoksille 95%:n luottamusvälit (Vidakovic 2011, 255) täydentämään intervention vaikutuksen suuruuden arviointia.

Cohenin (1988) mukaan efektikoko on pieni, kun Cohen's d=0,20, keskikoinen, kun d=0,50 ja suuri, kun d=0,80. Kun kyseessä on koe- ja kontrolliryhmien vertailu, d=0 tarkoittaa, että koe- ja kontrolliryhmien pistemäärien keskiarvo on sama. Jotta kaikkien (99,9%) kontrolliryhmäläisten pistemäärät olisivat pienempiä kuin koeryhmän keskiarvo, täytyy d:n olla 3,0 (Cohen 1988). Johdopäätöksiä tehtäessä yksittäistä havaittua efektikokoa verrataan aiemmin samasta aiheesta raportoituihin efektikokoihin (Thompson 2007, Lakens 2013).

Raja-arvoihin perustuvan efektikoon sanallisen kuvailun lisäksi efektikoolle voi laskea luottamusvälin, joka osaltaan konkreettisesti efektikoon tulkintaa. Iteoimalla tuotettu luottamusväli on efektikoon tapauksessa käyttökelpoisempi kuin kaavalla laskettu luottamusväli (Thompson 2007). 95%:n luottamusväli tarkoittaa, että satunnaisesti populaatiosta poimittu muuttujan arvo osuu

95%:n todennäköisyydellä luottamusvälin sisään. Otoksen perusteella laskettu luottamusväli voi sijoittua kokonaisuudessaan populaation keskiarvon ala- tai yläpuolelle (Thompson 2007).

### *Cohen's d efektikoko*

Kun Cohen's d efektikoko lasketaan otokselle, sen yleinen kaava on:

$$d_s = \frac{\bar{x}_2 - \bar{x}_1}{s} \quad (\text{Kaava 1})$$

missä d=Cohen' d,  $\bar{x}_1$  ja  $\bar{x}_2$ =keskenään verrattavien ryhmien keskiarvot, esimerkiksi testipistemäärien keskiarvo ennen ( $\bar{x}_1$ ) ja jälkeen ( $\bar{x}_2$ ) intervention ja s=verrattavien ryhmien yhdistetty otoskeskihajonta (Kaava 1, Lakens 2013). s:n yleinen kaava on:

$$s = \sqrt{\frac{s_1^2(n_1-1) + s_2^2(n_2-1)}{n_1 + n_2 - 2}} \quad (\text{Kaava 2})$$

missä  $s_1$  ja  $s_2$ =verrattavien ryhmien otoskeskihajonnat ja  $n_1$  ja  $n_2$ =verrattavien ryhmien havainto- eli tilastoyksiköiden (tyypillisesti osallistujien) lukumäärät (Kaava 2).

Cohen's d efektikoko voidaan laskea myös populaatiolle, jolloin sen yleinen kaava on:

$$d_{pop} = \frac{\bar{x}_2 - \bar{x}_1}{\sigma} \quad (\text{Kaava 3})$$

missä d=Cohen's d,  $\bar{x}_1$  ja  $\bar{x}_2$ =keskenään verrattavien ryhmien keskiarvot ja  $\sigma$ =verrattavien ryhmien yhdistetty keskihajonta (Kaava 3, Rosnow & Rosenthal 2003).  $\sigma$ :n yleinen kaava on:

$$\sigma = s \sqrt{\frac{df}{n_1 + n_2}} \quad (\text{Kaava 4})$$

missä s=vertailtavien ryhmien yhdistetty otoskeskihajonta (Kaava 2), df=vapausasteiden lukumäärä= $n_1 + n_2 - 2$  ja  $n_1$  ja

$n_2$  = verrattavien ryhmien tilastoyksiköiden lukumäärät (Kaava 4). Populaatio eli perusjoukko (N) tarkoittaa kaikkien havaintoyksiköiden muodostamaa kokonaisuutta. Koska N on harvoin tiedossa, lasketaan yleensä  $d_s$  eikä  $d_{pop}$ .

### Hedges' g efektikoko

Jos otoskoko on pieni ( $n < 20$ ), Cohen'  $d_s$  on harhainen estimaatti populaatioon yleistettävästä efektiosta. Korjattu efektikoko, Hedges' g, lasketaan kaavalla:

$$g = \left(1 - \frac{3}{4 \times (n_1 + n_2) - 9}\right) d \quad (\text{Kaava 5})$$

missä  $n_1$  ja  $n_2$  = verrattavien ryhmien tilastoyksiköiden lukumäärät ja  $d$  = Cohen's  $d_s$  (Kaava 5, Lakens 2013).

### Cohen's h efektikoko

Cohen's h efektiin yleinen kaava on:

$$h = 2 \times (\text{arkussini} \sqrt{p_1}) - 2 \times (\text{arkussini} \sqrt{p_2})$$

(Kaava 6)

missä  $p_1$  ja  $p_2$  = verrattavien ryhmien populaatio-osuudet (Kaava 6, Rosnow & Rosenthal 2003). Arkussini on sinifunktion käänteisfunktio. Tässä artikkelissa esitettävissä laskuissa kulman suuruuden yksikkönä käytetään radiaania, ei astetta.

Arkussini-muunnoksen luottamusvälin yleinen kaava on:

$$\left[ \sin^2 \left( \text{arkussini} \sqrt{p_i} - \frac{z_{1-\alpha/2}}{2\sqrt{n_i}} \right), \sin^2 \left( \text{arkussini} \sqrt{p_i} + \frac{z_{1-\alpha/2}}{2\sqrt{n_i}} \right) \right]$$

(Kaava 7)

missä  $z_{1-\alpha/2}$  on standardinormaalijakauman  $1-\alpha/2$  persentiili (Kaava 7, Vidakovic 2011,

255). 95%:n luottamusvälillä  $\alpha=0,05$  ja  $1-\alpha/2=1-0,025=0,9750$ , jolloin standardinormaalijakauman kertymäfunktion taulukosta nähdään, että persentiili on 1,960.

### Meta-analyysin tekeminen ja tulkinta

Meta-analyysi on osa järjestelmällistä prosessia, jonka tarkoituksena on tuottaa tutkimukseen perustuva mahdollisimman luotettava ja yleistettävä vastaus johonkin täsmällisesti määriteltyn kysymykseen. Prosessi etenee vaiheittain, alkaen tutkimuskysymyksen määrittelemisestä ja päättyen yhdistetyn efektiin laskemiseen ja tulosten raportointiin. Tutkimuskysymyksen määrittelemisen jälkeen ennen varsinaista meta-analyysia tehdään systemaattinen kirjallisuuskatsaus, jolla varmistetaan, että kaikki olennaisesti asiaankuuluvat laadukkaat tutkimukset tulevat mukaan analyysiin. (Shorten & Shorten 2013, Malmivaara & Komulainen 2014.)

Meta-analyysissä voi käyttää joko satunnaismallia tai kiinteää mallia (Hedges & Vevea 1998, Borenstein ym. 2009). Satunnaismalli soveltuu useampiin tilanteisiin, koska siinä analysoitavat tutkimukset oletetaan otokseksi populaatiosta, jossa yksittäiset efektikoot vaihtelevat niihin vaikuttavien suorien ja epäsuorien tekijöiden vaikutuksesta. Kiinteä malli olettaa, että on olemassa yksi ainoa oikea efektikoko, josta luottavimman arvion antavat suuret tutkimukset. Kiinteä malli soveltuu tilanteisiin, joissa vaikuttavat tekijät tunnetaan ja pystytään kontrolloimaan. Tässä artikkelissa raportoiduissa meta-analyysissä käytettiin satunnaismallia ja ne tehtiin Comprehensive Meta-Analysis -tietokoneohjelman 3. versiolla (Biostat, Englewood, NJ, USA).

Meta-analyysin tulokset esitetään tavallisesti metsikkökuviona. Metsikkökuvioista käyvät ilmi meta-analyysissä yhdistettyjen yksittäisten tutkimusten efektikoot, efektikokojen painotettu keskiarvo eli meta-analyysin tulos ja niiden luottamusvälit. Usein metsikkökuviossa esitetään myös kunkin alkupe-  
räistutkimuksen painokerroin eli sen prosentuaalinen vaikutus meta-analyysin tulokseen. Meta-analyysin ”hyvyyden” kannalta olennaista ei ole yhdistettävien tutkimusten lukumäärä (Valentine ym. 2010). Olennaisia seikkoja ovat tutkimusten homogeenisuus eli keskinäinen samankaltaisuus (Higgins 2008) ja tutkimusten tilastollinen kyky havaita se, mitä niiden on oletettu havaitsevan (Turner ym. 2013), esimerkiksi kliinisesti merkittäväksi arvioitu ero ryhmien välillä.

Perinteisesti tutkimusten efektikokojen homogeenisuuden testaamiseksi on laskettu niin kutsuttu Cochran’s Q (Kaava 8) (Hedges & Vevea 1998), jota kuitenkin on kritisoitu sen heikosta kyvystä todentaa heterogeenisuutta, jos tutkimuksia on vähän (Higgins ym. 2003).

$$Q = \sum_{i=1}^k W_i (Y_i - M)^2 \quad (\text{Kaava 8})$$

missä  $W_i$  on tutkimuksen  $i$  painokerroin,  $Y_i$  tutkimuksen  $i$  efektikoko,  $M$  efektikokojen ( $i-k$ ) keskiarvo ja  $k$  tutkimusten lukumäärä (Hedges & Vevea 1998, Borenstein ym. 2009).  $Q$  on siis efektikokojen painotettu neliösumma.

Cochran’s  $Q$  -tunnusluvun vaihtoehdoksi on ehdotettu  $I^2$ -tunnuslukua, joka ilmoittaa heterogeenisuuden prosenttilukuna välillä 0–100 (Higgins ym. 2003).  $I^2$  voidaan laskea helposti  $Q$ :n avulla (Kaava 9):

$$I^2 = \left( \frac{Q - df}{Q} \right) \times 100 \% \quad (\text{Kaava 9})$$

missä  $Q$ =Cochran’s  $Q$  (Kaava 8) ja  $df=k-1$ , kun  $k$ =tutkimusten lukumäärä.  $I^2$  on käytökelpoinen, koska se on asteikosta riippumaton ja siten vertailtavissa eri meta-analyysien kesken.  $I^2$  ilmoittaa todellisen heterogeenisuuden osuuden havaitusta vaihtelusta.  $I^2=0\%$  tarkoittaa, että tutkimusten välillä ei ole lainkaan aitoa vaihtelua ja  $I^2=100\%$  tarkoittaa, että kaikki havaittu vaihtelu on aitoa, ei siis otantavirheestä johtuvaa. Todellisen heterogeenisuuden osuuden sanotaan olevan pieni, kun  $I^2 \leq 25\%$ , kohtalainen, kun  $I^2 \leq 50\%$  ja suuri, kun  $I^2 \geq 75\%$ . (Higgins ym. 2003.)

Tutkimusten välistä heterogeenisuutta voi mitata myös  $T^2$ -tunnusluvulla, joka on havaittujen efektikokojen perusteella laskettu todellisten efektikokojen varianssin estimaatti (Kaava 10).

$$T^2 = \frac{Q - df}{C} \quad (\text{Kaava 10})$$

missä  $Q$ =Cochran’s  $Q$  (Kaava 8),  $df=k-1$ , kun  $k$ =tutkimusten lukumäärä, ja  $C$  on suure (Kaava 11), jolla  $Q$ :n arvot muutetaan takaisin alkuperäisten efektikokojen mitta-asteikolle ja keskiarvoistetaan.

$$C = \sum W_i - \frac{\sum W_i^2}{W_i} \quad (\text{Kaava 11})$$

missä  $W_i$  on tutkimuksen  $i$  painokerroin (Borenstein ym. 2009).

Vastaavasti  $T = \sqrt{T^2}$  on todellisten efekti-kokojen keskiarvon keskihajonnan estimaatti. Periaatteessa  $T$ :n avulla on mahdollista arvioida meta-analyttisesti yhdistettyjen yksittäisten tutkimusten efekti-kokojen keskiarvon ( $M$ ) vaihteluväli. Yleisesti oletetaan, että vaihteluväli  $\cong 2T - M + 2T$  (Borenstein ym. 2009). Koska heterogeenisuuden arvioinnissa käytettävät tunnusluvut mittaavat kukin hieman eri asiaa, meta-analyysistä ei kannata tehdä johtopäätöksiä pelkästään yhden tunnusluvun perusteella (Higgins 2008).

## Tulokset

### *Cohen's d efekti-koko*

Kaavoissa 14 ja 15 on esitetty Cohen's  $d$  efekti-koon laskeminen tutkimuksessa Tsai ym. (2008) raportoidulle interventiolle, jossa sairaanhoitajien osaamista liittyen keskuskokemokateetriin annettaviin lääkeinjektioihin pyrittiin parantamaan virtuaalisimulaatio-opetuksella. Intervention vaikutusta selvitettiin tietotestillä, joka koostui 10 kysymyksestä, joihin vastattiin joko kyllä tai ei. Cohen's  $d$  efekti-koon laskemista varten tutkimuksesta poimittiin seuraavat interventiota koskevat tiedot: tietotestiin ennen ja jälkeen intervention osallistuneiden lukumäärät ( $n_1 = 37$ ,  $n_2 = 37$ ), interventiota edeltäneessä testissä saatujen pistemäärien keskiarvo ja keskihajonta ( $\bar{x}_1 = 13,7$ ,  $s_1 = 2,0$ ) sekä intervention jälkeisessä testissä saatujen pistemäärien keskiarvo ja keskihajonta ( $\bar{x}_2 = 15,9$ ,  $s_2 = 1,6$ ). Ensin laskettiin  $s$ :

$$s = \sqrt{\frac{2,0^2(37-1) + 1,6^2(37-1)}{37+37-2}} = \sqrt{\frac{236,16}{73}} \cong \sqrt{3,235} \cong 1,799 \quad (\text{Kaava 12})$$

Sitten laskettiin  $\sigma$ : Siinä tarvittava  $s$  saatiin kaavasta 12.



$$\sigma = \sqrt{\frac{236,16}{73}} \times \sqrt{\frac{37+37-2}{37+37}} \cong 1,799 \times \sqrt{0,973} \cong 1,774 \quad (\text{Kaava 13})$$

Sitten laskettiin Cohen's d efektikoko otokselle:

$$d_s = \frac{15,9-13,7}{\sqrt{\frac{236,16}{73}}} \cong 1,223 \quad (\text{Kaava 14})$$

Ja lopuksi laskettiin Cohen's d populaatiolle:

$$d_{pop} = \frac{15,9-13,7}{1,774} \cong 1,240 \quad (\text{Kaava 15})$$

Tässä tapauksessa  $n \neq N$ , joten tuloksissa raportoitiin  $d_s$ .

#### *Hedges' g efektikoko*

Kaavassa 16 on esitetty Hedges' g efektiin laskeminen tutkimuksessa Tsai ym. (2008) raportoidulle interventiolle.  $d_s$  saatiin kaavasta 14.

$$g = \left(1 - \frac{3}{4 \times (37+37) - 9}\right) \left(\frac{15,9-13,7}{\sqrt{\frac{236,16}{73}}}\right) \cong 0,990 \times 1,223 \cong 1,211 \quad (\text{Kaava 16})$$

#### *Cohen's h efektikoko*

Tutkimuksessa Otero ym. (2008), jossa virheellisiä lääkitystapahtumia pyrittiin vähentämään sairaalan yleistä toimintakulttuuria parantamalla, ennen interventiota havainnoitiin yhteensä 1764 lääkitystapahtumaa, joista virheellisiä oli 201. Virheiden populaatio-osuus oli täten:

$$p_1 = \frac{K_2}{n_2} = \frac{201}{1764} \cong 0,114 \text{ eli } 11,4\% \quad (\text{Kaava 17})$$

Kaavassa 18 on esitetty Cohen's h efekti-  
koon laskeminen tutkimuksessa Otero ym.  
(2008) raportoidulle interventiolle:

$$h = 2 \times \left( \text{arkussini} \sqrt{\frac{201}{1764}} \right) - 2 \times \left( \text{arkussini} \sqrt{\frac{199}{2732}} \right) \cong 2 \times \left( \text{arkussini} \sqrt{0,114} \right) - 2 \times \left( \text{arkussini} \sqrt{0,073} \right) \cong 2 \times 0,344 - 2 \times 0,274 \cong 0,689 - 0,547 \cong 0,142$$

Kaavassa 19 on esitetty arkussini-muunnok-  
sen luottamusvälin laskeminen tutkimukses-  
sa Otero ym. (2008) raportoidulle ennen in-  
terventiota havaitulle populaatio-osuudelle:

$$\left[ \sin^2 \left( \text{arkussini} \sqrt{\frac{201}{1764} - \frac{1,960}{2\sqrt{1764}}} \right), \sin^2 \left( \text{arkussini} \sqrt{\frac{201}{1764} + \frac{1,960}{2\sqrt{1764}}} \right) \cong \sin^2(0,344 - 0,023), \sin^2(0,344 + 0,023) \cong \sin^2(0,321), \sin^2(0,367) \cong 0,100, 0,129 \right]$$

(Kaava 19)

Vastaavasti arkussini-muunnoksen luotta-  
musväli tutkimuksessa Otero ym. (2008) ra-  
portoidulle intervention jälkeen havaitulle  
populaatio-osuudelle  $\cong 0,064, 0,083$  eli 6,4–  
8,3%.

### *Meta-analyysi*

Taulukoissa 1 ja 2 on esitetty meta-ana-  
lyysien metsikkökuviot. Taulukossa 1 tutki-  
muksessa Tsai ym. (2008) raportoidun in-  
tervention efekti-koot (Cohen's d ja Hedges'  
g) eroavat hieman kaavoissa 14 ja 16 laske-  
tuista arvoista, koska käsin laskussa käytet-  
tiin likiarvoja ja tietokoneohjelma käytti tark-  
koja arvoja. Tutkimusten, joissa interventi-  
on tulokset oli raportoitu testipistemäärinä,  
meta-analyysin  $I^2$  oli 31,7%, mikä tarkoittaa,  
että aineistossa oli kohtalaisesti aitoa hete-  
rogeenisuutta (vrt. Higgins ym. 2003). Saman  
meta-analyysin  $T^2$  oli 2,10, mikä voidaan tul-  
kita esimerkiksi siten, että yksittäisen tutki-  
muksen efekti-koon (Hedges' g) poikkeama  
efektikokojen keskiarvosta oli keskimäärin  
 $\sqrt{2,1}=1,45$ . Tutkimusten, joissa tulokset oli

raportoitu prosentiosuuksina, meta-analyy-  
sin  $I^2$  oli 17,0% eli aineistossa oli vähän aitoa  
heterogeenisuutta (vrt. Higgins ym. 2003).

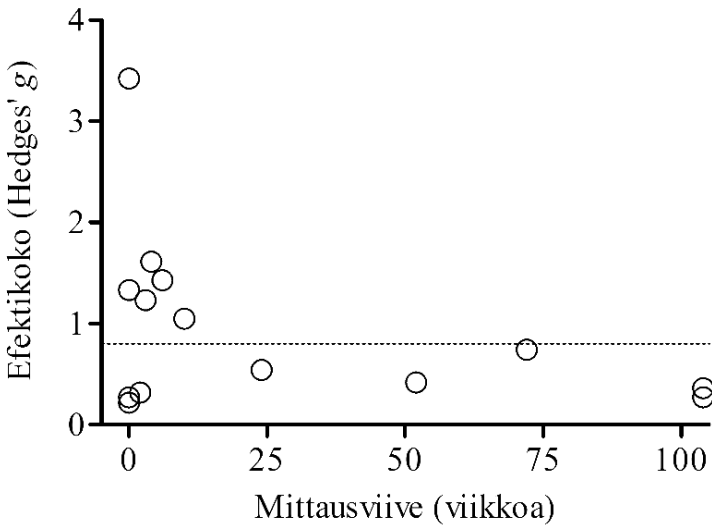
### **Pohdinta**

#### *Tulosten tarkastelu*

Tutkimuksissa, joissa intervention vaiku-  
tusta selvitettiin taito- tai tietotestillä, efek-  
tikoko (Hedges' g) oli  $>1$  viidessä tapauk-  
sessa seitsemästä (Taulukko 1). Vastaavasti  
tutkimuksissa, joissa intervention vaikutus-  
ta selvitettiin havainnoimalla aitoja lääkitys-  
tapahtumia ja raportoimalla niissä tapahtu-  
neet virheet, efekti-koko oli kaikissa tapauk-  
sissa  $<0,7$  (Taulukko 2). Edellä esitetyn pe-

rusteella ei voi kuitenkaan päätellä, että ”testitutkimuksissa” interventioiden vaikutus oli keskimäärin selvästi suurempi kuin ”havainnointitutkimuksissa”, sillä testitutkimuksissa viive interventiosta sen arviointiin oli lyhempi kuin havainnointitutkimuksissa. Intervention vaikutus oli poikkeuksetta  $<0,8$  (Hedges'  $g$ ), kun viive intervention ja sen vaiku-

tuksen mittaamisen välillä oli yli 10 viikkoa (Kuvio 1). Lisäksi voidaan ajatella, että suhteellinen muutos lääkitysvirheiden määrässä on intervention vaikuttavuuden mittana parempi kuin muutos sairaanhoitajien taitotai tietotasossa, joka itsessään ei vielä tarkoita laadukkaampaa hoitoa.



Kuvio 1. Intervention vaikutuksen (Hedges'  $g$  efektikoko) subde intervention ja sen vaikutuksen mittaamisen väliseen aikaviiveeseen. Katkoviiva osoittaa suuren efektikoon raja-arvon (0,8).

Tämän tutkimuksen aineiston muodostivat 14 alkuperäistutkimusta, joissa yhdessäkään intervention vaikutusta ei ollut arvioitu efektikokoja laskemalla. Koska kaikki 14 tutkimusta olivat sangen tuoreita ja julkaistu vertaisarvioituissa kansainvälisissä tiedelehdissä, on interventioiden vaikutuksen arvioinnin kehittymättömyys laaja ja ajankohdainen ongelma hoitotieteessä. Intervention vaikutuksesta tai vaikuttavuudesta ei voi tehdä johtopäätöksiä ainoastaan intervention tilastollisen merkitsevyyden perusteella. Jos intervention vaikutus todennetaan mittaamalla osaamista, jonka kohentamiseen interventio on suunniteltu ja vertaamalla mittaustulosta ei-interventiota -tilanteeseen, on epäloogista olettaa, että interventiolla ei

olisi haluttua vaikutusta (ks. esim. Cook ym. 2008, 2010). Ja koska tilastollisessa testauksessa nollahypoteesina on nimenomaan vaihtoehto, että mitattavassa ominaisuudessa ei tapahdu muutosta, saa vastahypoteesi lähestulkoon aina tukea.

Intervention vaikutus on ainoastaan yksi sitä huomattavasti laajemman vaikuttavuuden käsitteen osatekijöistä (Barreto 2005). Hyvästäkään interventiosta ei tule vaikuttavaa, jos sen implementointi eli käyttöönotto epäonnistuu. Implementaatio on monitekijäinen prosessi, jonka jokaisella osa-alueella on omat tunnusomaiset haasteensa (ks. esim. Damschroder ym. 2009). Yksi haasteista on intervention käyttökelpoisuuden todentaminen siinä nimenomaisessa kon-

tekstissä, johon sitä ollaan implementoimassa. Erityisesti hoitotyössä, missä implementoitavat interventiot ovat yleensä käytännönläheisiä, intervention vaikutuksen ja vaikuttavuuden konkretisoiminen on implementoinnin onnistumisen kannalta ensiarvoisen tärkeää. Toinen keskeinen haaste on intervention kohteiden ja/tai käytännön toteuttajien eli useimmiten hoitohenkilökunnan havaintojen, kokemusten ja mielipiteiden riittävä ja oikeanlainen huomioiminen (Sidani ym. 2016). Jos intervention kohteet ja/tai toteuttajat mieltävät intervention huonoksi, implementointi tuskin onnistuu, vaikka interventio olisi numeroiden avulla todistettu vaikuttavaksi.

### *Tutkimuksen luotettavuus*

Tässä tutkimuksessa käytetty aineisto koostui keskenään varsin erilaisista interventioista, mikä ei ole ihanteellinen lähtökohta meta-analyysille. Toisaalta  $I^2$ -tunnusluvun perusteella aidon heterogeenisuuden osuus oli meta-analyysissä pieni tai kohtalainen, minkä voidaan katsoa puoltavan meta-analyysin tekemistä. Meta-analyysin luotettavuuden arvioinnissa on tärkeää erottaa toisistaan tunnusluvuilla ilmaistava vastemuuttujan vaihteluun perustuva numeerinen heterogeenisuus ja yhdistettyjen tutkimusten toiminnallinen johdonmukaisuus. Toiminnallisesti epäjohdonmukaisen meta-analyysin tuloksesta ei voi tehdä vahvoja johtopäätöksiä ja yleistyksiä, vaikka tunnuslukujen perusteella laskennallinen heterogeenisuus ei olisi ongelma. Meta-analyysin tuloksen luotettavuutta arvioitaessa on siis vastattava kysymykseen: Ovatko yhdistettävät tutkimukset keskenään riittävän samantaisia, jotta niiden voidaan sanoa kuvaavan samaa tarkasti määriteltyä ilmiötä? Tässä artikkelissa esitettyjen kaavojen ja laskujen luotettavuus on varmistettu käyttämällä useita lähteitä ja laskemalla laskut sekä käsin että tietokoneohjelmalla.

## **JOHTOPÄÄTÖKSET JA SUOSITUKSET**

Tätä menetelmäartikkelia voi käyttää sikirijana efektikokoja laskettaessa ja meta-analyysin tuloksia tulkittaessa. Esitellyt efektikoot, Cohen's d, Cohen's h ja Hedges' g, ovat kaikkiin tilanteisiin soveltuvia vaikutuksen suuruuden mittoja, mutta eivät suinkaan ainoita mahdollisia. Cohenin ja Hedgesin efektikoot ovat käyttökelpoisia erityisesti, kun on tarpeen muuntaa useita erilaisia vaikutuksen tunnuslukuja "yleispäteväksi" efektikooksi ja siten mahdollistaa tutkimusten yhdistäminen meta-analyytisesti.

Tässä aineistossa intervention vaikutuksen suuruus oli yhteydessä intervention toteutuksen ja sen arvioinnin väliseen aikaväiveeseen. Havainto tähdentää intervention vaikutuksen ja vaikuttavuuden erottamista toisistaan. Vaikutus liittyy intervention välittömään vaikutukseen ja vaikuttavuus intervention varsinaiseen lopputulemaan, jonka mittaaminen tyypillisesti edellyttää pitempikestoista seurantaa. Lisäksi havainto korostaa yhdistettävien tutkimusten toiminnallisen johdonmukaisuuden arviointia meta-analyysin tuloksia tulkittaessa. Vaikka tutkimuksissa olisi käytetty samaa vastemuuttujaa ja siihen olisi pyritty vaikuttamaan samanlaisella interventiolla, voi meta-analyysin johdonmukaisuuden kyseenalaistaa, mikäli viive interventiosta sen arviointiin vaihtelee suuresti tutkimusten välillä.

Ilmiöiden syvällisen ymmärtämisen ja hoitotyön käytäntöjen kehittämisen kannalta on ensiarvoisen tärkeää suunnitella interventioita, joiden testaaminen puolestaan edellyttää kokeellista tutkimusta. Interventoiden implementoinnissa voidaan keskittyä toimivimpiin ja tehokkaimpiin interventioihin, kun niiden vaikutusta arvioidaan efektikokojen ja meta-analyysien avulla. Hoitotieteen tutkijoiden tulisi vahvistaa tutkimuksen laatua hakemalla korkeinta mahdollista näyttöä interventioiden vaikutuksesta ja vaikuttavuudesta.

## VASTUUALUEET

Tutkimuksen suunnittelu: AV, MH, aineistonkeruu: MH, aineiston analysointi: AV, kä-

sikirjoituksen kirjoittaminen: AV, käsikirjoituksen kommentointi: MH

## LÄHTEET

- Abbasinazari M., Zareh-Toranposhti S., Hassani A., Sistanizad M., Azizian H. & Panahi Y. (2012) The effect of information provision on reduction of errors in intravenous drug preparation and administration by nurses in ICU and surgical wards. *Acta Medica Iranica* **50**(11), 771–777.
- Barreto M.L. (2005) Efficacy, effectiveness, and the evaluation of public health interventions. *Journal of Epidemiology & Community Health* **59**(5), 345–346.
- Blank F.S., Tobin J., Macomber S., Jaouen M., Dinoia M. & Visintainer P. (2011) A “back to basics” approach to reduce ED medication errors. *Journal of Emergency Nursing* **37**(2), 141–147.
- Borenstein M., Hedges L.V., Higgins J.P.T. & Rothstein H.R. (2009) *Introduction to Meta-Analysis*. John Wiley & Sons, Chichester, UK.
- Chamberlain B. (2007) The difference between meta-analysis and meta-study: part I. *Clinical Nurse Specialist* **21**(5), 229–230.
- Cohen J. (1988) *Statistical Power Analysis for the Behavioral Sciences*. 2. edition. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Cook D.A., Levinson A.J., Garside S., Dupras D.M., Erwin P.J. & Montori V.M. (2008) Internet-based learning in the health professions. *The Journal of the American Medical Association* **300**(10), 1181–1196.
- Cook D.A., Erwin P.J. & Triola M.M. (2010) Computerized virtual patients in health professions education: a systematic review and meta-analysis. *Academic Medicine* **85**(10), 1589–1602.
- Damschroder L.J., Aron D.C., Keith R.E., Kirsh S.R., Alexander J.A. & Lowery J.C. (2009) Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation Science* **4**(50), 1–15.
- EPHPP. (2016) Quality Assessment Tool for Quantitative Studies. <http://www.ephpp.ca/tools.html> (20.10.2016)
- Fineout-Overholt E. & Johnston L. (2005) Teaching EBP: asking searchable, answerable clinical questions. *Worldviews on Evidence-Based Nursing* **2**(3), 157–160.
- Ford D.G., Seybert A.L., Smithburger P.L., Kobulinsky L.R., Samosky J.T. & Kane-Gill S.L. (2010) Impact of simulation-based learning on medication error rates in critically ill patients. *Intensive Care Medicine* **36**(9), 1526–1531.
- Harne-Britner S., Kreamer C.L., Frownfelter P., Helmuth A., Lutter S., Schafer D.J. & Wilson C. (2006) Improving medication calculation skills of practicing nurses and senior nursing students: a pilot study. *Journal for Nurses in Staff Development* **22**(4), 190–195.
- Hedges L.V. & Vevea J.L. (1998) Fixed- and random-effects models in meta-analysis. *Psychological Methods* **3**(4), 486–504.
- Higgins J.P.T. (2008) Commentary: heterogeneity in meta-analysis should be expected and appropriately quantified. *International Journal of Epidemiology* **37**(5), 1158–1160.
- Higgins J.P.T., Green S. (editors). (2011) *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. [www.handbook.cochrane.org](http://www.handbook.cochrane.org). (20.10.2016)
- Higgins J.P.T., Thompson S.G., Deeks J.J. & Altman D.G. (2003) Measuring inconsistency in meta-analyses. *BMJ* **327**(7414), 557–560.
- Hohenhaus S., Cadwell S., Stone-Griffith S., Sears-Russell N., Baxter T., Hicks W., Maples L. & Kleja K. (2008) Assessment of emergency nursing practice during critical pediatric medication administration in a simulated resuscitation using the “Color Coding Kids Hospital System”. *Advanced Emergency Nursing* **30**(3), 233–241.
- HOTUS. (2016) Hoitotyön tutkimussäätiö. Näyttöön perustuva toiminta. Vaikuttavuutta terveydenhuoltoon. Tutkimusten kriittisen arvioinnin kriteeristö. <http://www.hotus.fi/jbi-fi/kriittinen-arviointi> (20.10.2016)
- Härkänen M., Voutilainen A., Turunen E. & Vehviläinen-Julkunen K. (2016) Systematic review and meta-analysis of educational interventions designed to improve medication administration skills and safety of registered nurses. *Nurse Education Today* **41**, 36–43.
- Kliger J., Blegen M.A., Gootee D. & O'Neil E. (2009) Empowering frontline nurses: a structured intervention enables nurses to improve medication administration accuracy. *Joint Commission Journal of Quality and Patient Safety* **35**(12), 604–612.
- Korhonen A., Hakulinen-Viitanen T., Jylhä V. & Holopainen A. (2013) Meta-synthesis and evidence-based health care – a method for systematic review. *Scandinavian Journal of Caring Sciences* **27**(4), 1027–1034.
- Lakens D. (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology* **4**(863), 1–12.
- Lu M.C., Yu S., Chen I.J., Wang K.W., Wu H.F. & Tang F.I. (2013) Nurses' knowledge of high-alert medications: a randomized controlled trial. *Nurse Education Today* **33**(1), 24–30.
- Malmivaara A. & Komulainen J. (2014) Luotettavaa vaikuttavuustietoa järjestelmällisistä katsauksista. *Duodecim* **130**(16), 1635–1641.

- Otero P., Leyton A., Mariani G., Ceriani Cernadas J.M. & Patient Safety Committee. (2008) Medication errors in pediatric inpatients: prevalence and results of a prevention program. *Pediatrics* **122**(3), e737–e743.
- Rosnow R.L. & Rosenthal R. (2003) Effect sizes for experimental psychologists. *Canadian Journal of Experimental Psychology* **57**(3), 221–237.
- Schmidt F.L. (1992) What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist* **47**(10), 1173–1181.
- Schneider P.J., Pedersen C.A., Montanya K.R., Curran C.R., Harpe S.E., Bohenek W., Perratto B., Swaim T.J. & Wellman K.E. (2006) Improving the safety of medication administration using an interactive CD-ROM program. *American Journal of Health-System Pharmacy* **63**(1), 59–64.
- Sherriff K., Burston S. & Wallis M. (2012) Effectiveness of a computer based medication calculation education and testing programme for nurses. *Nurse Education Today* **32**(1), 46–51.
- Shorten A. & Shorten B. (2013) What is meta-analysis? *Evidence Based Nursing* **16**(1), 3–4.
- Sidani S., Manojlovich M., Doran D., Fox M., Covell C.L., Kelly H., Jeffs L. & McAllister M. (2016) Nurses' perceptions of interventions for the management of patient-oriented outcomes: a key factor for evidence-based practice. *Worldviews on Evidence-Based Nursing* **13**(1), 66–74.
- Simonsen B.O., Daehlin G.K., Johansson I. & Farup P.G. (2014) Improvement of drug dose calculations by classroom teaching or e-learning: a randomised controlled trial in nurses. *BMJ Open* **4**(10), e006025.
- Sullivan G.M. & Feinn R. (2012) Using effect size – or why the P value is not enough. *Journal of Graduate Medical Education* **4**(3), 279–282.
- Sung Y.H., Kwon I.G. & Ryu E. (2008) Blended learning on medication administration for new nurses: integration of e-learning and face-to-face instruction in the classroom. *Nurse Education Today* **28**(8), 943–952.
- Thompson B. (2007) Effect sizes, confidence intervals, and confidence intervals for effect sizes. *Psychology in the Schools* **44**(5), 423–432.
- Tsai S.L., Chai S.K., Hsieh L.F., Lin S., Taur F.M., Sung W.H. & Doong J.L. (2008) The use of virtual reality computer simulation in learning Port-A cath injection. *Advances in Health Science Education: Theory and Practice* **13**(1), 71–87.
- Turner R.M., Bird S.M. & Higgins J.P.T. (2013) The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PLoS One* **8**(3), e59202.
- Valentine J.C., Pigott T.D. & Rothstein H.R. (2010) How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics* **35**(2), 215–247.
- Vidakovic B. (2011) *Statistics for Bioengineering Sciences: With MATLAB and WinBUGS Support*. Springer, New York, NY.
- Xu C., Li G., Ye N. & Lu Y. (2014) An intervention to improve inpatient medication management: a before and after study. *Journal of Nursing Management* **22**(3), 286–294.
- Zeng X., Zhang Y., Kwong J.S., Zhang C., Li S., Sun F., Niu Y. & Du L. (2015) The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. *Journal of Evidence-Based Medicine* **8**(1), 2–10.

*Ari Voutilainen, FT, dosentti, projektitutkija, Itä-Suomen yliopisto, Hoitotieteen laitos, Yliopistonranta 1C, PL 1627, 70211 Kuopio, ari.voutilainen@uef.fi*

*Marja Härkänen, TtT, tutkijatohtori, Itä-Suomen yliopisto, Hoitotieteen laitos, Yliopistonranta 1C, PL 1627, 70211 Kuopio, marja.harkanen@uef.fi*