

JBI: Arviointikriteerit diagnostisen testin tarkkuustutkimukselle

Tätä tarkistuslistaa käytetään diagnostisen testin tarkkuustutkimuksen metodologisen laadun arviointiin. Arvioinnin tarkistuslistaan sisältyy yhteensä 10 arviointikriteeriä, joiden yksityiskohtaiset sisällöt on lyhyesti kuvattu alla. Arvioijan on hyvä tutustua myös JBI:n julkaisemaan katsauksen tekijöiden käsikirjaan, jonka pdf-linkki löytyy Tutkimusten arviointikriteeristö -kotisivulta arviointia tehdessään. Tarkistuslistan alkuperäinen englanninkielinen versio löytyy kotisivumme kohdasta alkuperäiset tutkimusten arviointikriteeristöt. Kunkin kriteerin toteutuminen arvioidaan asteikolla: Kyllä (K), Ei (E), Epäselvä (?), Ei sovellettavissa (NA). (Whiting ym. 2011, Campbell ym. 2015.)

Arvioija _____ Päiväys _____

Tekijä(t) _____ Vuosi _____ Nro _____

Potilaan valikointi

K E ? NA

1. Käytettiin peräkkäis- tai satunnaisotantaa? ☐ ☐ ☐ ☐

2. Vältettiin tapaus-verrokkitutkimusasetelmaa? ☐ ☐ ☐ ☐

3. Vältettiin tarpeettomia tutkittavien poissulkemisia? ☐ ☐ ☐ ☐

Indeksitesti

K E ? NA

4. Tulkittiinko indeksitestin tulokset ilman tietoa vertailutestin tuloksista? ☐ ☐ ☐ ☐

5. Jos diagnostisia raja-arvoja käytettiin, niin olivatko ne ennalta määritetty? ☐ ☐ ☐ ☐

6. Luokitteiko vertailutesti kiinnostuksen kohteena olevan sairauden/tilan todennäköisesti oikein? ☐ ☐ ☐ ☐

7. Tulkittiinko vertailutestin tuloksia ilman tietoa indeksitestin tuloksista? ☐ ☐ ☐ ☐

8. Oliko indeksi- ja vertailutestin välillä sopiva aikaväli? ☐ ☐ ☐ ☐

9. Tehtiinkö kaikille potilaille sama vertailutesti? ☐ ☐ ☐ ☐

10. Sisällytettiin kaikki potilaat analyysiin? ☐ ☐ ☐ ☐

Kokonaisarviointi: Hyväksy ☐ Hylkää ☐ Lisätietoja tarvitaan ☐

Kommentteja (mukaan lukien syy hylkäykseen):

Kriittisen arvioinnin tarkistuslista diagnostisen testin tarkkuustutkimukselle

Lähteet: Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM & QUADAS-2 Group (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 155(8): 529-36.

Campbell JM, Klugar M, Ding S, Carmody DP, Hakonsen SJ, Jadotte YT, White S & Munn Z (2015) Diagnostic test accuracy: methods for systematic review and meta-analysis. *International Journal of Evidence-Based Healthcare* 13(3): 154-62.

Potilaan valikointi

1. Käytettiin peräkkäis- tai satunnaisotantaa?

Tutkimuksessa pitää olla mainittu ja kuvattu, miten tutkittavat valittiin tutkimukseen. Jos tutkimuksessa on valittu satunnaisotos, niin satunnaisotannan menetelmä tulee ilmoittaa (ja sen tulisi olla soveltuva). Peräkkäisotantaa käytettäessä tulisi kuvata muutakin kuin, että "n-määrä peräkkäisiä tutkittavia otettiin mukaan", eli olisi kuvattava myös miten peräkkäiset tutkittavat valittiin. Esimerkiksi "Tutkimukseen otettiin mukaan kaikki potilaat (n-määrä), joilla on osteosarkooma, ja joita hoidettiin yksikössämme maaliskuun 2018 ja kesäkuun 2018 välillä".

Arvioi: Mitä otantamenetelmää tutkimuksessa käytettiin? Kuvattiinko otantamenetelmä riittävällä tarkkuudella ja oliko menetelmä soveltuva?

2. Vältettiin tapaus-verrokkitutkimusasetelmaa?

Tapaus-verrokkitutkimuksia on kuvattu tarkemmin Joanna Briggs Instituutin julkaisemassa katsauksen tekijöiden käsikirjassa. Lyhyesti: Jos tutkimukseen rekrytoidaan tutkittavia, joilla on jo jollakin toisella tavalla diagnosoitu tutkittava sairaus tai tila, ja sen jälkeen arvioidaan tutkittavalla testillä tunnistaako kyseinen testi näiden potilaiden sairastavan kyseistä sairautta tai tilaa, niin vastaus tähän kysymykseen on "ei".

Arvioi: Mikä on tutkimuksen asetelma? Vältettiin tapaus-verrokkiasetelmaa?

3. Vältettiin tarpeettomia tutkittavien poissulkemisia?

Jos tutkittavia jätetään pois sellaisista syistä, jotka todennäköisesti vaikuttavat testin suorittamiseen, tulkintaan tai testin tuloksiin, tämä voi vääristää tuloksia. Esimerkkejä tarpeettomasta tutkittavien poissulkemisesta: Tutkittava jätetään pois tutkimuksesta, koska

- hänelle on vaikea tehdä testiä
- tutkittavan tulos on raja-arvolla
- tutkittavalla on selkeä oirekuva kiinnostuksen kohteena olevasta diagnoosista.

Arvioi: Millä perusteilla tutkittavia suljettiin tutkimuksesta pois? Vältettiin tarpeettomia poissulkukriteereitä?

Indeksitesti

4. Tulkittiinko indeksitestin tulokset ilman tietoa vertailutestin tuloksista?

Indeksitestin tulosten tulkitsejan pitäisi olla sokkoutettu vertailutestin tuloksille. Jos vertailutestiä ei ole tehty siinä vaiheessa, kun indeksitesti tehdään ja tulkitaan, niin vastaus tähän kysymykseen on ”kyllä”. Jos sama ihminen tulkitsee indeksitestin ja vertailutestin, niin tähän kysymykseen vastataan oletettavasti ”ei”, mikäli testiin ei liity muita sokkouttavia tekijöitä. Sokkouttavana tekijänä voi olla esimerkiksi se, että testi tehdään ja analysoidaan/tulkitaan eri aikaan, ja että testin tulkitseja voidaan sokkouttaa potilaiden identiteetiltä ja vertailutestien tuloksilta testin tulkitavaiheessa.

Arvioi: Oliko indeksitestin tulkinta sokkoutettu vertailutestin tuloksista? Eli se, joka tulkitsee indeksitestin tuloksen, ei ollut tietoinen kyseisen tutkitavan vertailutestin tuloksesta.

5. Jos diagnostisia raja-arvoja käytettiin, niin olivatko ne ennalta määritetty?

Diagnostisia raja-arvoja voidaan määrittää datasta sen mukaan, mitkä arvot antavat optimaalisen tarkkuuden testille tai ne voidaan määrittää ennalta. Jos diagnostisia raja-arvoja ei ole käytetty (ts. testitulokset perustuu tietyn ominaisuuden havaitsemiseen, mikä joko on olemassa tai ei ole olemassa), tähän kysymykseen vastataan NA.

Arvioi: Olivatko diagnostiset raja-arvot (jos sellaisia käytettiin) ennalta määritetty vai määritettiin ne tutkimuksen datasta?

6. Luokitteiko vertailutesti kiinnostuksen kohteena olevan sairauden/tilan todennäköisesti oikein?

Vertailutestin pitäisi olla yleisesti parhaana pidetty menetelmä (golden standard) kiinnostuksen kohteena olevan tilan tai sairauden diagnosointiin. Lisäksi tutkimuksessa tulisi raportoida testin suorittaminen riittävän yksityiskohtaisesti, jotta lukija voi olla varma siitä, että testi on tehty ja sitä on käytetty oikein.

Arvioi: Käytettiinkö vertailutestinä yleisesti parhaana pidettyä menetelmää kiinnostuksen kohteena olevan tilan/sairauden diagnosointiin? Tehtiinkö ja käytettiinkö testiä oikein?

7. Tulkittiinko vertailutestin tuloksia ilman tietoa indeksitestin tuloksista?

Kriteerin 4 selostus koskee myös tätä kohtaa. Vertailutestin tulokset pitäisi tulkita joku, joka on sokkoutettu indeksitestin tuloksilta. Jos indeksitestiä ei ole tehty siinä vaiheessa, kun vertailutesti tehdään ja tulkitaan, niin vastaus tähän kysymykseen on ”kyllä”. Jos sama ihminen tulkitsee vertailutestin ja indeksitestin, niin tähän kysymykseen vastataan oletettavasti ”ei”, mikäli testiin ei liity muita sokkouttavia tekijöitä. Sokkouttavana tekijänä voi olla esimerkiksi se, että testi tehdään ja analysoidaan/tulkitaan eri aikaan, ja että testin tulkitseja voidaan sokkouttaa potilaiden identiteetiltä ja vertailutestien tuloksilta testin tulkitavaiheessa.

Arvioi: Oliko vertailutestin tulkinta sokkoutettu indeksitestin tuloksista? Eli se, joka tulkitsee vertailutestin tuloksen, ei ollut tietoinen kyseisen tutkittavan indeksitestin tuloksesta.

8. Oliko indeksi- ja vertailutestin välillä sopiva aikaväli?

Indeksitesti ja vertailutesti tulisi suorittaa ajallisesti riittävän lähellä toisiaan, jotta potilaan tilanne testausten välillä ei olisi muuttunut. Suurin sallittu testausten välinen aika vaihtelee kiinnostuksen kohteena olevan populaation ja testin ominaisuuksien mukaan.

Arvioi: Toteutettiin indeksi- ja vertailutesti ajallisesti riittävän lähekkäin, jotta potilaan tilanne ei olisi oleellisesti muuttunut testausten välillä?

9. Tehtiinkö kaikille potilaille sama vertailutesti?

Vertailutestin, jonka mukaan potilaat luokitellaan (heillä on tai ei ole kiinnostuksen kohteena olevaa sairautta tai tilaa), pitäisi olla sama kaikille potilaille. Jos indeksitestin tulokset vaikuttavat siihen, miten vertailutestiä käytetään tai käytetäänkö sitä (esimerkiksi, jos ilmeinen väärä negatiivinen tulos havaitaan, niin tehdään "kaksoistarkastus"), tämä voi vinouttaa testin herkkyyden ja spesifisyyden arviointia. Lisäksi joissakin tutkimuksissa voidaan käyttää kahta rinnakkaista vertailutestiä (eri potilailla), joiden tulokset yhdistetään. Kummassakin tapauksessa tähän kysymykseen vastataan "ei".

Arvioi: Tehtiinkö kaikille potilaille sama vertailutesti?

10. Sisällytettiin kaikki potilaat analyysiin?

Tutkittavien kato tutkimuksen aikana pitäisi raportoida. Kadon syitä ja laajuutta tulisi pohtia, ja sitä onko niillä mahdollisesti ollut vaikutuksia tuloksiin (tähän kysymykseen vastatessa voi käyttää subjektiivista harkintaa, esimerkiksi ylipäättään alhainen tutkittavien pysyvyys tutkimuksessa ohjaa vastaamaan "ei", mutta joidenkin yksittäisten tutkittavien poisjäänti suuresta kohortista ei välttämättä tarkoita negatiivista vastausta tähän kysymykseen). Jos tutkittavan testitulosten vaikea tulkittavuus aiheuttaa hänen poisjättämisensä analyysistä, voi tulokset liioitella diagnostisen testin tarkkuusarviota, ja tähän kysymykseen tulisi ehdottomasti vastata "ei".

Arvioi: Raportoitiinko kadon syyt ja laajuus ja pohdittiinko niiden vaikutuksia tuloksiin? Sisällytettiin kaikki tutkittavat analyysiin?