

# Julkaisuharha – lääketieteellisen tiedon akilleenkantapä

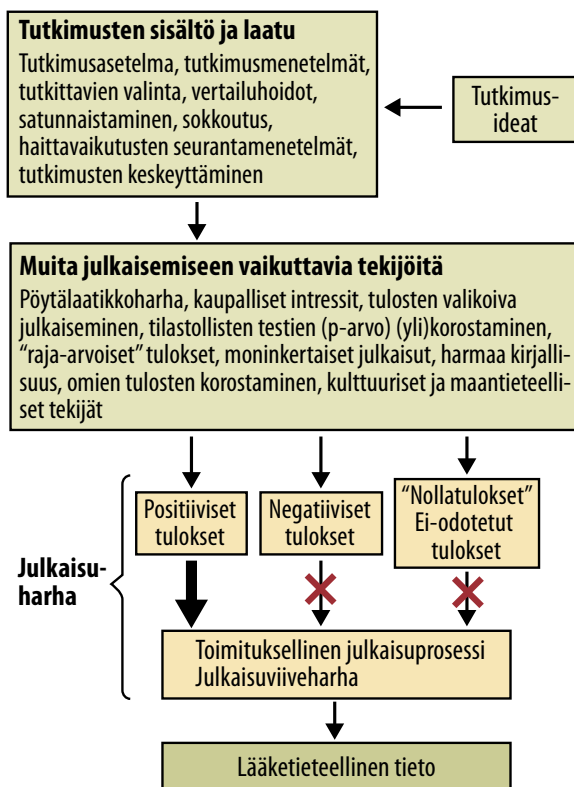
Tiedeyhteisön taipumusta julkaista herkemmin myönteisiä tai odotettuja kuin kielteisiä tai odottamattomia tuloksia kutsutaan julkaisuharhaksi. Se vaikuttaa lääketieteellisen tiedon kertymiseen monella tavalla. Positiivisten tai odotuksia vastaavien tulosten yliarviointi voi johtaa tiedon vääristymiseen ja vaikuttaa hoitokäytäntöihin, jotka perustuvat puutteellisiin meta-analyysihin ja katsauksiin. Cochrane-katsauksessa on arvioitu, että positiivisten tulosten julkaiseminen on enimmillään neljä kertaa todennäköisempää kuin negatiivisten. Kielteisten tutkimustulosten julkaisuviive on jopa muutama vuoden pidempi kuin myönteisten tulosten. Myös muut tekijät kuin julkaisuharha vaikuttavat julkaisujen valikoitumiseen. Näitä ovat pöytälaattikoharha, harmaa kirjallisuus, tulosten valikoiva julkaiseminen sekä kaupalliset intressit tai niiden puute. Julkaisuharhaa voidaan vähentää rekisteröimällä alkavat tutkimukset tietokantoihin, sopimalla tutkimuksen aineiston käyttöoikeudesta, suosimalla avointa vertaisarviointia, käyttämällä julkaisuharhaa arvioivia tilastollisia menetelmiä meta-analyysissä ja vähentämällä tilastollisen merkitsevyyden ylikorostamista.

**Näyttöön perustuvan lääketieteen** aikakaudella näytön tulisi olla luotettavaa, ei puutteellista tai harhaista. Julkaisuharhalla tarkoitetaan sitä, että myönteisiä tai odotuksia vastaavia tutkimustuloksia julkaistaan useammin kuin muunlaisia tuloksia (Olson ym. 2002). Odotusten vastaiset tulokset voivat olla ”nollatuloksia” (ryhmien välillä ei ole tilastollisesti

merkitseviä eroja), kielteisiä (vertailuryhmässä tulokset ovat alkuoletusten vastaisesti parempia kuin koeryhmässä) tai muuten ei-toivottuja. Tällaisia ovat esimerkiksi tulokset, joilla ei ole lainkaan kliinistä merkitystä tai merkitys on hyvin vähäinen (KUVA). Lääketieteellisen tutkimuksen julkaisuharha on ”akateeminen tauti”, jolla on useampi kuin yksi ilmenemismuoto ja jonka patogeneesi on monimutkainen. Osittain kyse on myös sensuurista, koska odotuksia vastaavien tulosten yliarviointi voi johtaa tiedon vääristymiseen ja vaikuttaa siten hoitokäytäntöihin.

Lääketieteellisissä lehdissä julkaisuharhaa on pohdittu jo viime vuosisadan alussa. Vuonna 1909 Boston Medical Journalin pääkirjoituksessa kuvattiin menestyksellisten tapauselostusten valikoivaa siteerausta (Dickersin 1990). Suomalaisessa lääketieteellisessä kirjallisuudessa kliinisten lääketutkimusten julkaisemisharhaa on pohdittu melko vähän (Hemminki 1980, Keränen ja Ylitalo 1999). Julkaisuharha voi syntyä tieteellisen tiedon käsitteilyprosessin eri vaiheissa, ennen julkaisuhetkeä ja sen jälkeen. Tutkijat vaikuttavat itse julkaisuiden laatuun ja tulkintaan, mutta myös lehtien toimittajien ja vertaisarvioijien rooli on huomattava, kun he valikoivat tuloksia julkaistavaksi (Chalmers ym. 1990). Jos tutkimus on julkaistu, se voi päästä mukaan systemoituun katsaukseen tai meta-analyysiin, joita pidetään näytönasteeltaan luotettavimpina tiedon muodostamisen tapoina. Meta-analyysien tulosten vaikutukset kliiniseen päätöksentekoon ja terveyspolitiikkaan ovat merkittävät. Koska meta-analyysissä ei voida huomioida julkaisematta, arvioimatta tai kirjoittamatta jätettyjä tutkimustuloksia, päätelmät saattavat olla puutteellisia tai jopa harhaanjohtavia.





**KUVA.** Julkaisuharha ja muita julkaisujen laatuun, sisältöön ja toteutumiseen vaikuttavia tekijöitä lääketieteellisen tiedon muodostumisessa.

Muiden julkaisujen valikoitumiseen vaikuttavien tekijöiden pohtiminen on tavanomaisen julkaisuharhan huomioon ottamisen lisäksi tärkeää, koska molemmista voi olla vakavia seurauksia kliiniselle päätöksenteolle. Tämän katsauksen tarkoituksena on näiden seikkojen lisäksi arvioida julkaisuharhan suuruutta lääketieteen eri osa-alueilla kliinisen tutkimuksen näkökulmasta sekä pohtia julkaisuharhan vähentämisen keinoja.

### Julkaisuharhan suuruus

PubMed-tietokannasta löytyi 3 710 julkaisua hakusanalla ”publication bias” (22.7.2011). Kun haku rajattiin niihin julkaisuihin, joiden otsikossa oli ”publication bias”, jäljelle jäi 260 artikkelia. Artikkeleista 251 oli englanninkielisiä. Pääkirjoituksia oli 31, katsauksia 35 ja me-

ta-analyysijä 16. Tästä kirjoituksesta rajattiin pois yhdeksän muun kuin englanninkielistä artikkelia, eikä hakuja tehty muista tietokannoista, mikä voi aiheuttaa harhaa myös tähän katsaukseen. Toisaalta PubMed sisältää suuren osan lääketieteellisestä tutkimuksesta, joten valikointi- ja kieliharhan suuruus on todennäköisesti tässä katsauksessa vähäinen.

**TAULUKKON** on koottu erikoisaloittain ne PubMedistä löytyneet julkaisut, joissa on pyritty arvioimaan julkaisuharhan suuruutta määrällisesti käyttäen lähteenä erikoisalan vuosittaisia tieteellisiä kokouksia ja niiden abstrakteja, tutkimuseettisiä tai muita tutkimusrekistereitä ja tietokantoja sekä MEDLINE- tai PubMed-hakuja. Julkaisuharhan suuruudeksi arvioitiin enimmillään jopa 50 %. Ortopedian alalla julkaisuharhan suuruudeksi arvioitiin 12–19 %. Ainoastaan 35 %:ssa ortopedisistä

**TAULUKKO.** Julkaisuharhan suuruutta arvioivia tutkimuksia 2000-luvulla erikoisaloittain.

Erikoisala	Lähde	Aineisto	Tulokset
Anestesiologia	Hall ym. 2007	T	44 % julkaistu <sup>1</sup> , 85 %:ssa julkaistuista positiivinen löydös Julkaisuajan mediaani 32,5 kk Ensimmäisen ja toisen vaiheen lääkekokeita julkaistiin vähemmän kuin kolman- nen tai neljännen vaiheen
Ortopedia	Harris ym. 2007, Vavken ja Dorotka 2011	A	55 % julkaistu <sup>1</sup> Julkaisuharhan suuruus 12–19 % Positiiviset (OR 1,62, 95 %:n lv 1,01–2,6) tai merkitsevät tulokset (OR 2,05, 95 %:n lv 1,2–3,4) todennäköisempiä Meta-analyseistä 35 % arvioi julkaisuharhaa
Neurologia (aivohalvaus)	Liebeskind ym. 2006	R, T	Julkaisemattomissa (4 kpl) enemmän haitallisia tuloksia (75 % vs 6 %, p < 0,0001) kuin julkaistuissa Tutkimuksissa, joissa ei osoitettu hyötyjä hoidosta, pidempi julkaisuviive
Neurokirurgia	Yoshimoto 2003	T	Takautuvissa tutkimuksissa ei-rupturoituneiden aneurysmien sairastavuus + kuol- leisuus 7,8 %, monikeskustutkimuksissa 20,3 %, meta-analyyseissä 5–12,7 %
Infektio-epi- demiologia	O'Brien ym. 2006	R	55 % julkaistu <sup>1</sup>
Akuuttihoito	Ospina ym. 2006	A	51 % julkaistu <sup>1</sup> Julkaisu aika keskimäärin 32 kk, 59 % RCT:istä julkaistu viiden vuoden kuluessa Julkaistuissa käsikirjoituksissa tulokset tukivat harvemmin hypoteesia kuin abstrakteissa
Gastroentero- logia	Harewood 2005	T	EUS:n tarkkuus tuumorin koon suhteen raportoitu 40 tutkimuksessa, imu- solmukemetastasoinnin suhteen 27:ssä EUS:n kyky peräsuolisöyvän asteen arvioinnissa voi olla yliarvioitu
Gastroentero- logia	Timmer ym. 2002	A	47 % julkaistu <sup>1</sup> Hyväksytyllä abstraktilla 2,3-kertainen (95 %:n lv 1,7–3,1) todennäköisyys tulla julkaisuksi. Suuren impactiluvun lehdissä julkaisu on vähäisempää, jos tulokset eivät ole merkitseviä (OR 0,5, 95 %:n lv 0,3–0,6)
Lastentaudit	Klassen ym. 2002	A	59 % julkaistu <sup>1</sup> 64 % julkaistuista RCT:istä tuki uutta hoitoa, 44 % julkaisemattomista Keskimääräinen otoskoko oli julkaistuissa suurempi kuin julkaisematta jääneissä Julkaistuista tutkimuksista 5 %:ssa päätelmä muuttui abstraktiin verrattuna
Obstetriikka ja gynekologia	Evers 2000, Polyzos ym. 2011	A	56–57 % julkaistu <sup>1</sup> Julkaisut keskimäärin 15–32,5 kk:n kuluessa Vuonna 2000 arvioitu julkaisuharha vähäinen Vuonna 2011 positiiviset tulokset vaikuttivat (nopeampaan) julkaisuun
Perinatologia	Blackwell ym. 2009	A	56 % julkaistu <sup>1</sup> Positiivisia löydöksiä sisältäneet julkaistu useammin (60 % vs 33 %)
Psykiatria (masennus- lääkkeet)	Turner ym. 2008	R, T	Positiivisia tuloksista sisältäneistä julkaistu 99 %, negatiivisia tuloksia sisältäneis- tä 39 % 36:sta negatiivisen tuloksen saaneesta tutkimuksesta 11 oli julkaistu siten, että päätelmissä oli positiivinen löydös Kaikista julkaistuista tutkimuksista 94:ssä oli positiiviset tulokset; vain puolessa kaikista tutkimuksista oli positiiviset tulokset
Hematologia (luuydinsiirrot)	Saeed ym. 2011	A	30 % julkaistu <sup>1</sup> Positiivisia tuloksia julkaistu yhtä usein kuin negatiivisia eli julkaisuharhan todennäköisyys pieni Kliiniset tutkimukset julkaistu useammin kuin ei-kliiniset

<sup>1</sup>Julkaistujen käsikirjoitusten osuus tarjotuista.

R = rekisteri, A = abstrakti, T = tietokanta, RCT = satunnaistettu kontrolloitu tutkimus, EUS = endoskooppinen kaiku-  
kuvaus

meta-analyyseistä julkaisuharha oli huomioitu (Vavken ja Dorotka 2011). Jos abstrakti oli hyväksytty esitettäväksi tieteellisessä kokouksessa, todennäköisyys sen julkaisemiselle oli yli kaksinkertainen (Timmer ym. 2002). Akuutti-hoidon ja lastentautien osalta havaittiin julkaisutujen tulosten eroavan abstrakteissa mainituista tuloksista (Klassen ym. 2002, Ospina ym. 2006). Embryologian vuosikokousten abstrakteista julkaisuharhaa ei havaittu vielä vuonna 2000, mutta vuonna 2011 se todettiin selvästi (Evers 2000, Polyzos ym. 2011).

Aivoinfarktia koskevassa katsauksessa huomattiin, että julkaisemattomissa tutkimuksissa oli enemmän haittoja kuvaavia tuloksia kuin julkaistuissa (Liebeskind ym. 2006). Sama tulos saatiin masennuslääketutkimuksista. Näitä valmisteita koskevista myönteisiä tuloksia saaneista tutkimuksista oli julkaistu lähes kaikki (37/38) mutta vain 39 % (14/36) tutkimuksista, joissa löydökset olivat kielteisiä (Turner ym. 2008). Niistäkin useimmissa (11/14) päätelmäksi oli nostettu positiivinen löydös. Luuydinsiirtoja koskeva katsaus oli ainoa, jossa ei havaittu eroa, joten julkaisuharhan todennäköisyys arvioitiin pieneksi (Saeed ym. 2011).

Tutkimukset, joissa ei saada merkitseviä tuloksia, julkaistaan keskimäärin myöhemmin kuin tutkimukset, joiden tulokset ovat merkitseviä (viiveharha, time lag bias) (Scherer ym. 2007). Neurologiassa keskimääräinen viive negatiivisten tutkimustulosten julkaisemisessa on yli kaksi vuotta (Liebeskind ym. 2006) (TAULUKKO).

Julkaisuharhan suuruutta koskevissa laajoissa katsauksissa on saatu erikoisalojen yksittäiskysymyksiä suurempia arvioita. Archives of Internal Medicine -lehdessä vuonna 2008 arvioiduista 100 käsitteistä 77 %:ssa oli positiivisia tuloksia. Myönteisiä tuloksia sisältävistä käsitteistä 83 % julkaistiin lehdessä mutta vain 5 % niistä käsitteistä, joiden tulokset olivat kielteisiä. Hopewell ym. (2009) arvioivat Cochrane-tietokannasta julkaisuharhaa yhdistetyllä riskisuhteella ja saivat tulokseksi lähes nelinkertaisen todennäköisyyden positiivisten tulosten julkaisemiselle negatiivisiin verrattuna (OR 3,90, 95 %:n lv 2,68–5,68). Tämä lienee suurin esitetty arvio julkaisuharhan laajuudesta, joten ongelman mittavuutta on vaikea kiistää.

### Julkaisujen valikoitumiseen liittyviä muita tekijöitä

Tavanomaisen myönteisten ja kielteisten tulosten valikoitumisen lisäksi julkaisuharhaa muodostavat tulosten kirjoittamatta jättäminen (file drawer bias, pöytälaatikkoharha), harmaa kirjallisuus (grey literature) eli opinnäytteet tai muut raportit, tulosten valikoiva julkaiseminen, monien samansuuntaisia tuloksia sisältävien julkaisujen tekeminen ja tulosten julkaisematta jättäminen kaupallisten intressien tai niiden puutteen vuoksi. Muita julkaisujen tuloksiin vaikuttavia tekijöitä ovat muun muassa tutkimuksen rakenne, tutkittavien valinta, satunnaistaminen, sokkoutus, vertailuhoidot,

### JULKAISUHARHAA VÄHENTÄVIÄ KEINOJA

- ▶ Kokeellisten tutkimusten rekisteröinti tietokantaan
- ▶ Aineiston käyttöikeudesta sopiminen ennen tutkimuksen alkua (tutkimussopimukset)
- ▶ Lehtien vertaisarvioinnin käytäntöjen laadun parantaminen, muun muassa avoin vertaisarviointi
- ▶ CONSORT- ja STROBE-ohjeistusten noudattaminen tulosten raportoinnissa
- ▶ Julkaisuharhaa arvioivat ja korjaavat tilastotieteelliset menetelmät meta-analyyseissä
- ▶ Näytön arviointi sekä kokeellisten että ei-kokeellisten tutkimusten perusteella
- ▶ Tilastollisen merkitsevyyden (p-arvon) ylikorostamisesta luopuminen
- ▶ Luottamusvälien hyödyntäminen tutkimuksen tulosten tulkinnassa

haittojen seurantamenetelmät, tilastokäsittely, tutkimuksen aloittamatta jättäminen tai keskeyttäminen ja omien tulosten korostaminen (KUVA). Nämä eivät kuitenkaan sisälly julkaisuharhaan, joten tilastollista käsittelyä lukuun ottamatta niitä ei käsitellä tässä katsauksessa enempää.

Tutkijoiden aika ei aina riitä havaintojen raportointiin kaikesta kerätystä aineistosta. Tulosten kirjoittamatta jättäminen voi toki olla harkittuakin, koska tutkimusrahoituksessa saatetaan suosia uusien aineistojen keruuta jo kerättyjen aineistojen analyysin sijaan. Harmaata kirjallisuutta ei löydy tietokannoista, ja siksi se voi jäädä myös meta-analyysien ulkopuolelle (Burdett ym. 2003).

Valikoivasta julkaisemisesta on käytetty nimitystä publication bias in situ (Phillips 2004). Sillä viitataan tarkoituksellisesti in situ -syöpään, sillä kyseessä on ilmiö, joka on sidoksissa vain tiettyyn erikoisalaan ja kysymykseen, eikä sitä voi ennakoita tai systemaattisesti estää. Tutkijat ovat tärkeässä asemassa, kun he valitsevat tutkimuksensa menetelmiä ja kysymyksenasetteluja tai kun he päättävät tulosten analyysistä ja niiden julkaisemisesta osissa tai kokonaisuudessaan. Negatiivisiin tuloksiin johtavia seikkoja ovat tutkimuskysymysten ja erityisesti mitattavien vastemuuttujien valinta – nämäkin ovat molemmat tutkijoiden päätöksiä. Kielteisen tuloksen saaminen on mahdollista, jos vastemuuttujaksi on valittu epäherkkä muuttuja, esimerkiksi kehon painon muutos. Painoon vaikuttavat lihasmassan määrä, rasvakudoksen suhteellinen osuus, kehon nestepitoisuus ja monet muut perinnölliset tekijät. Laihdutukseen tähtäävässä tutkimuksessa voidaan päätyä nollatulokseen, koska kokonaispaino on summamuuttuja. Valittaessa vastemuuttujaksi esimerkiksi sisäelinten ympärillä oleva rasva, voidaan saada isojakin eroja ryhmien välille, vaikka painossa ryhmien välillä ei olisikaan eroa. Tutkimusaineiston keruu- ja analyysivaiheet ovat pitkiä, ja uusia vastemuuttujiin liittyviä tuloksia todennäköisesti julkaistaan prosessin aikana. Dwan ym. (2008) totesivat katsauksessaan, että 40–62 %:ssa tutkimuksista vähintään yksi pääasiallisista tutkimusohjelmassa mainituista

vastemuuttujista oli muutettu tai vaihdettu tai muuttujaa koskeva tulos oli jätetty raportoitamatta.

Aina ei ole mahdollista saada käyttöön alkuperäisiä kliinisten tutkimusten tuloksia, etenkin jos tutkimusten kustannuksista vastaa kaupallinen tahon. Esimerkiksi haavanhoito negatiivisen paineen avulla (negative pressure wound therapy) on laajalti käytössä, vaikka näyttö menetelmän hyödyistä on vähäistä eikä julkaisuharhan suuruutta ole onnistuttu arvioimaan tietojen rajoittamisen takia (Peinemann ym. 2008).

Julkaisuharhaan vaikuttavat monet tekijät sekä suoraan että välillisesti (Song ym. 2009). Julkaisuprosessiin osallistuvat henkilöt tutkijoista toimittajiin ovat suorassa yhteydessä julkaisuharhaan. Epäsuorasti julkaisuharhaan vaikuttavat lehtien monet positiiviset tutkimuslöydökset ja pienten tutkimusten suuret vaikutusten suuruudet (effect size). Näiden lisäksi myös kulttuuriset ja maantieteelliset tekijät saattavat vaikuttaa julkaisuharhaan. Vaihtoehtoisia hoitoja (CAM, complementary and alternative medicine) koskevat tutkimukset olivat eurooppalaisissa suuren impakttiluvun lääketieteellisissä lehdissä useammin tuloksiltaan positiivisia verrattuna niihin tutkimuksiin, jotka julkaistiin yhdysvaltalaisissa lehdissä (Sood ym. 2007).

## Julkaisuharhan vähentäminen

Lääketieteen informaattikot arvioivat 2000-luvun alussa, että sähköiset lehdet ovat ratkaisu julkaisuharhaan, kun tilaa on enemmän ja julkaisukynnystä saataisiin näin alemmaksi (Song ym. 1999, Friedman ja Wyatt 2001). Suomalaiset julkaisuharhaa pohtineet asiantuntijat toivoivat 2000-luvun alussa internetin lisääntyvän käytön ja tutkimusrekisterien yleistymisen vähentävän julkaisuharhaa (Keränen ja Ylitalo 1999). Tieteelliset lehdet ovat sittemmin siirtyneet lähes kokonaan sähköisiksi, mutta julkaisuharha ei silti ole poistunut, vaikka se

**Positiivinen tutkimustulos tulee julkaistuksi jopa neljä kertaa todennäköisemmin kuin negatiivinen tulos**

onkin ehkä jonkin verran vähentynyt. Suurin julkaisuharhaa vähentävä toimenpide on 2000-luvulla ollut tieteellisen näytön läpinäkyvyyden lisääntyminen. Keinoina tässä ovat olleet kliinisten kokeiden jo olemassa olevat ja ei-kokeellisille tutkimuksillekin ehdotetut tutkimusrekisterit, joiden avulla voidaan jäljittää tutkimusten alkuperäisiä kysymyksenasetteluja ja tavoitteita suhteessa julkaistuihin tietoihin (Protocol registration system, [www.clinicaltrials.gov](http://www.clinicaltrials.gov)). Kliinisten kokeiden ja havainnoivien tutkimuksen raportointiohjeet (CONSORT, CONSolidated Standards of Reporting Trials, STROBE, Strengthening The Reporting of Observational studies in Epidemiology) (Loder ym. 2010) ja muut ohjeistukset ovat saattaneet lisätä tutkimusten raportoinnin laatua ainakin tasokkaimmissa lehdissä.

Tieteellisissä lehdissä julkaisuharhaa on pyritty vähentämään myös muun muassa perustamalla kielteisten tulosten osioita (esim. Journal of Cerebral Blood and Flow Metabolism, Journal of Clinical Endocrinology and Metabolism, Cancer Epidemiology Biomarkers Prev) (Dirnagl ja Lauritzen 2010). PLoS (Public Library of Open Science) on kaupallista hyötyä tavoittelematon taho, jonka perustamissa lehdissä on pyritty tietoisesti vähentämään kliinisten kokeiden julkaisuharhaa. Vuonna 2005 perustettu PLoS Clinical Trials on verkkolehti tai arkisto, johon hyväksytään kliinisiä kokeita riippumatta saaduista tuloksista, aineiston koosta tai tulosten merkityvyydestä (Veitch 2005). British Medical Journal on julkaissut negatiivisten tulosten erikoisnumeron vuoden 2012 alussa masennuslääkkeisiin liittyvästä julkaisuharhasta käynnistyneen keskustelun ansiosta (Loder 2011).

Meta-analyysien julkaisuharhan arvioimiseen ja korjaamiseen on useita menetelmällisiä keinoja (Poorolajal ym. 2010, Guyatt ym. 2011). Menetelmissä vertaillaan katsaukseen kuuluvien tutkimusten tulosten epäsymmetriaa suhteessa otoskokoon (mm. funnel plot). Erityisesti systemoiduissa katsauksissa 2000-luvulla on yhä useammin pyritty julkaisuharhan huomioimiseen. Englantilaisen tutkimuksen mukaan viidesosassa vuonna 2006 tehdyistä systemoiduista katsauksista oli et-

sitty muita kuin englanninkielisiä tutkimuksia, 61 %:ssa oli haettu myös julkaisemattomien tutkimusten tuloksia ja neljäsosassa käytettiin tilastollisia menetelmiä julkaisuharhan arvioimiseksi (Parekh-Bhurke ym. 2011). Hollantilaiset selvittivät, onko oikeutettua jättää julkaisematta kokeellisia tutkimuksia, joiden voima jää alle 80 %:n (Borm ym. 2009). Tutkijat havaitsivat, että vain 60 % niistä kokeellisista tutkimuksista, joiden tilastollinen voima oli alle 50 %, julkaistaan, vaikka niiden tyyppin 1 virheen suuruus oli enimmillään vain 5 %. Voimalaskelmat eivät näin ollen yksinomaan oikeuta julkaisematta jättämiseen.

Tieteellisen läpinäkyvyyden lisäämisessä yhtenä tekijänä on vertaisarvioinnin laadun parantaminen avoimella refereeprosessilla. Tärkeä julkaisuharhaa vähentävä tekijä on olemassa olevan näytön kokonaisarviointi (totality of evidence), olipa sitten kyse kokeellisesta tai ei-kokeellisesta tutkimuksesta, tilastollisesti merkitsevistä tai ei-merkitsevistä havainnoista. Pelkkä tilastollisen merkityvyyden metsästys johtaa harhaan. Sen sijaan on ehdotettu käytettäväksi ennen tutkimusta tehtävää arviota väärien positiivisten havaintojen saamisesta (FPRP, false positive reporting probability) (Ioannidis ja Trikalinos 2007). FPRP riippuu sekä otoksen suuruudesta että ennen tutkimusta arvioidusta vaikutuksen suuruudesta (prior probability). Tämän vuoksi suurinkaan tutkimus ei pelkän otoskoon perusteella saavuta täysin virheetöntä tulosta, jos oletettu yhteys ennen tutkimuksen suorittamista on pieni. Pieni otoskoko, pieni havaittu vaikutus, suuri vertailujen lukumäärä, suuri diagnostisten kriteerien vaihtuvuus ja lukumäärä sekä suuret taloudelliset tai muut vastaavat intressit lisäävät todennäköisyyttä saada (väärä) positiivisia tuloksia (Ioannidis ja Trikalinos 2007). Tilastolliselta merkityvyydeltään ”raja-arvoille” päätyvät tulokset (borderline significance) ovat olleet tarkastelun kohteena hiljattain. Raja-arvoisten tulosten saaminen ei välttämättä tarkoita sitä, että vaikutusta ei löytyisi: p-arvon raja 0,05 on vain yhteisesti sovittu raja, mutta sitä pidetään virheellisesti eksaktina osoituksena vaikuttavuudesta. Luottamusvälit ovat paras tapa ar-

vioida tuloksen merkitystä – todellinen tulos on todennäköisesti luottamusvälin keskikohdassa (Hackshaw ja Kirkwood 2011, *Duodecim* 1/1988, biostatistiikan teemanumero).

## Lopuksi

Lääketiede on käytännönläheistä asiantuntijatyötä, jota ohjaavat lääketieteellisen tutkimuksen näyttöön perustuvat arviot sairauksien ehkäisystä ja hoidosta. Näyttöön perustuvalla lääketieteelläkin on heikkoutensa. Kyse on inhimillisestä myönteisten asioiden suosimisesta tiedon tuottamisen ja sen hyödyntämisen prosesseissa. Francis Bacon kirjoitti vuonna 1605 ilmestyneessä kirjassaan *The advancement of learning* ihmislouheen olevan taipuvainen hyväksymään positiiviset havainnot useammin kuin negatiiviset. Bacon on epäsuorasti nimetty julkaisuharhan isähahmoksi (Petticrew 1998).

Julkaisuharhaa tietämättään tuottavat tutkijat ja lehtien toimittajat ja vertaisarvioijat toiminevat hyvien tieteellisten käytäntöjen mukaan tarkoituksenaan edistää oikeaa tietoa. ”Oikein” suoritettujen (kokeellisten) tutkimusten uskotaan tuottavan oletettuja ja toivottuja tuloksia. Tähän oletukseen eivät hyvin suoritettut mutta kielteisiin tuloksiin päätyvät tutkimukset sovi. Jos tutkimus päättyy nollatulokseen tai negatiivisiin tuloksiin, saatetaan olettaa tutkijoiden tehneen virheitä; he ovat valinneet väärän tutkimusasetelman, toteuttaneet sen puutteellisesti tai vähintään analysoineet tai tulkinneet tulokset väärin (Friedman ja Wyatt 2001). Tutkimusten tarkoitus voi olla joko hypoteesilähtöinen tai havainnoiva jolloin ei välttämättä olekaan tarkoitus päätyä yksiselitteiseen tulokseen. Tätä seikkaa

voidaan erityisesti tarkentaa, jotta myös vertaisarvioijat osaavat suhteuttaa tulokset tutkimuksen alkuperäiseen tarkoitukseen. Joskus tutkimuksessa voi tulla ilmi ”sivulöydöksiä”, joihin ei luoteta, koska niitä ei ryhdytty alun perin tutkimaan. Tieteellinen tutkimus ei kuitenkaan aina ole tuloksiltaan ennakoitua, joten tiukka tilastotieteellinen tai kliininen arviointi ei pelkästään riitä – tarvitaan tulosten kokonaisarviointia ja merkityksen tulkintaa.

*Archives of Internal Medicine* -lehden päätoimittajat ehdottavat, että lääketieteen eetoksen ”pyri välttämään epäsuotuisia vaikutuksia” (to do no harm) tulisi lisätä tutkijoiden kiinnostusta negatiivisten tutkimustulosten julkaisemiseen, jotta lääketieteessä voitaisiin tehdä kokonaisarvioita harhaisten positiivisten löydösten sijaan (Sridharan ja Greenland 2009). Asenteiden muuttuminen ja eetoksen toteuttaminen edellyttävät kielteisiä tuloksia sisältävien tutkimusten uudelleen määrittelyä. Friedman ja Wyatt (2001) luonnehtivat, että hyvä negatiivisen tuloksen sisältävä tutkimus tuottaa tärkeää tietoa ja hyötyä tuleville tutkimuksille ja käytännön päätöksille. Uuteen määritelmään voisi lisätä sivulauseen ”vaikka tilastollista merkitysvyyttä ei olisikaan osoitettu”. Näin päätelmät voitaisiin suhteuttaa tutkimuskysymyksen kokonaisperspektiiviin. ■

\* \* \*

Kiitokset professori Anssi Auviselle käsikirjoituksen kommentoinnista.

**RIITTA LUOTO, LT, dosentti, tutkimusjohtaja**  
UKK-instituutti ja  
Terveystieteiden tutkimuskeskus

**SIDONNAISUUDET**  
Kirjoittajalla ei ole sidonnaisuuksia

## Summary

### Publication bias – the Achilles’ heel of medical research

Overestimation of results that are positive or meet the expectations may lead to biased information and have an influence on treatment practices. Compared with positive research results, the publication delay of negative results may be longer by up to several years. The file drawer effect, grey literature, selective publication of results along with commercial interests or lack thereof also influence the selection of publications. Publication bias can be reduced by registering studies in databases, by agreeing upon rights to the results and by favoring open peer review.

## KIRJALLISUUTTA

- Blackwell SC, Thompson L, Refuerzo J. Full publication of clinical trials presented at a national maternal-fetal medicine meeting: is there a publication bias? *Am J Perinatol* 2009;26:679–82.
- Borm GF, den Heijer M, Zielhuis GA. Publication bias was not a good reason to discourage trials with low power. *J Clin Epidemiol* 2009;62:47.e1–10.
- Burdett S, Stewart LA, Tierney JF. Publication bias and meta-analyses: a practical example. *Int J Technol Assess Health Care* 2003;19:129–34.
- Chalmers TC, Frank CS, Reitman D. Minimizing the three stages of publication bias. *JAMA* 1990;263:1392–5.
- Dickersin K. The existence of publication bias and risk factors for its occurrence. *JAMA* 1990;263:1385–9.
- Dirnagl U, Lauritzen M. Fighting publication bias: introducing the negative results section. *J Cereb Blood Flow Metab* 2010;30:1263–4.
- Dwan K, Altman DG, Arnaiz JA, ym. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE* 2008;3:e3081.
- Evers JL. Publication bias in reproductive research. *Hum Reprod* 2000;15:2063–6.
- Friedwan CP, Wyatt JC. Publication bias in medical informatics. *J Am Med Inform Assoc* 2001;8:189–92.
- Guyatt GH, Oxman AD, Montori V, ym. GRADE guidelines: 5. Rating the quality of evidence- publication bias. *J Clin Epidemiol* 2011;64:1277–82.
- Hackshaw A, Kirkwood A. Interpreting and reporting clinical trials with results of borderline significance. *Br Med J* 2011; 343:d3340.
- Hall R, de Antueno C, Webber A; Canadian Research Ethics Board. Publication bias in medical literature: a review by a Canadian Research Ethics Board. *Can J Anaesth* 2007;54:380–8.
- Harewood GC. Assessment of publication bias in the reporting of EUS performance in staging rectal cancer. *Am J Gastroenterol* 2005;100:808–16.
- Harris IA, Mourad M, Kadir A, Solomon MJ, Young JM. Publication bias in abstracts presented to the annual meeting of the American Academy of Orthopaedic Surgeons. *J Orthop Surg (Hong Kong)* 2007;15:62–6.
- Hemminki E. Study of information submitted by drug companies to licensing authorities. *Br Med J* 1980;280:833–6.
- Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database of Systematic Reviews*. 2009;1:Art. No.:MR000006. DOI:10.1002/14651858.MR000006.pub3.
- Ioannidis JPA, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. *CMAJ* 2007;176:1091–6.
- Keränen T, Ylitalo P. Kliinisten lääketutkimuksen julkaisuharha. *Duodecim* 1999; 115:1828–32.
- Klassen TP, Wiebe N, Russell K, ym. Abstracts of randomized controlled trials presented at the society for pediatric research meeting: and example of publication bias. *Arch Pediatr Adolesc Med* 2002;156:474–9.
- Liebeskind DS, Kidwell CS, Sayre JW, Saver JL. Evidence of publication bias in reporting acute stroke clinical trials. *Neurology* 2006;67:973–9.
- Loder E. A theme issue in 2011 on unpublished evidence. *BMJ* 2011;342:d2627.
- Loder E, Groves T, MacAuley D. Registration of observational studies: The next step towards research transparency. *BMJ* 2010;340:c950.
- O'Brien SJ, Gillespie IA, Sivanesan MA, Elson R, Hughes C, Adak GK. Publication bias in foodborne outbreaks of infectious intestinal disease and its implications for evidence-based food policy. *England and Wales 1992-2003. Epidemiol Infect* 2006;134:667–74.
- Olson CM, Rennie D, Cook D, ym. Publication bias in editorial decision making. *JAMA* 2002;287:2825–8.
- Ospina MB, Kelly K, Klassen TP, Rowe BH. Publication bias of randomized controlled trials in emergency medicine. *Acad Emerg Med* 2006;13:102–8.
- Parekh-Bhurke S, Kwok CS, Pang C, ym. Uptake of methods to deal with publication bias in systematic reviews has increased over time, but there is still much scope for improvement. *J Clin Epidemiol* 2011;64:349–57.
- Peinemann F, McGauran N, Sauerland S, Lange S. Negative pressure wound therapy: potential publication bias caused by lack of access to unpublished study results data. *BMC Med Res Methodol* 2008;8:4.
- Petticrew M. Diagoras of Melos (500 BC): an early analyst of publication bias. *Lancet* 1998;352:1558.
- Phillips CV. Publication bias in situ. *BMC Med Res Methodol* 2004;4:20.
- Polyzos NP, Valachis A, Patavoukas E, ym. Publication bias in reproductive medicine: from the European Society of Human Reproduction and Embryology annual meeting to publication. *Hum Reprod* 2011;26: 1371–6.
- Poorolajal J, Haghdost AA, Mahmoodi M, Majdzadeh R, Nasserri-Moghaddam S, Fotouhi A. Capture-recapture method for accessing publication bias. *J Res Med Sci* 2010;15:107–15.
- Saeed M, Paulson K, Lambert P, Szwajcer D, Seftel M. Publication bias in blood and marrow transplantation. *Biol Blood Marrow Transplant* 2011;17:930–4.
- Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews* 2007; Issue 2. Art No; MR000005.
- Song F, Eastwood A, Gilbody S, Duley L. The role of electronic journals in reducing publication bias. *Med Inform Internet Med* 1999;24:223–9.
- Song F, Parekh-Burke S, Hooper L, ym. Extent of publication bias in different categories of research cohorts: a meta-analysis of empirical studies. *BMC Med Res Methodol* 2009;9:76.
- Sood A, Knudsen K, Sood R, ym. Publication bias for CAM trials in the highest impact factor medicine journals is partly due to geographical bias. *J Clinical Epidemiol* 2007;60:1123–6.
- Sridharan L, Greenland P. Editorial policies and publication bias: The importance of negative studies. *Arch Intern Med* 2009;169:1022–3.
- Timmer A, Hilsden RJ, Cole J, Hailey D, Sutherland LR. Publication bias in gastroenterological research – a retrospective cohort study based on abstracts submitted to a scientific meeting. *BMC Med Res Methodol* 2002;2:7.
- Turner EH. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;358: 252–60.
- Vavken P, Dorotka R. The prevalence and effect of publication bias in orthopaedic meta-analyses. *J Orthop Sci* 2011;16:238–44.
- Veitch E, PLoS Medicine Editors. Tackling publication bias in clinical trial reporting. *PLoS Medicine* 2005;2:e367.
- Yoshimoto Y. Publication bias in neurosurgery: lessons from series of unruptured aneurysms. *Acta Neurochir (Wien)* 2003;145:45–8.