



Hyvän artikkelin anatomia

3. osa: Tilastotiede

Etenkin uran alkuvaiheessa olevan lääkärin tai lääketieteen opiskelijan suurimmaksi huoleksi näyttää artikkelia arvioitaessa nousevan käytettyjen tilastollisten menetelmien arviointi. Tämä huoli voi kuitenkin olla liioiteltua, eikä tutkimuksen tulosten merkittävyyden pohtiminen aina vaadi kovinkaan syvällistä tilastollisten menetelmien tuntemusta. Oleellista on ymmärtää, mitä on tutkittu ja ovatko tutkimuksen tulokset luotettavia ja onko niillä merkitystä. Tilastotieteen käsitteiden ja menetelmien tulee olla lukijalle helposti ymmärrettäviä, ja käytännön lääkärin tulee tuntea keskeisimmät termit ja käsitteet.

Mielestämme artikkelin tulosten tilastollisessa arvioinnissa on myös käytettävä tervettä järkeä, ja usein jos tutkittava asia on tosi, se näkyy jo paljaalla silmällä. Kenties tilastotiedettä ei silloin edes tarvittaisi, mutta tulosten luotettavuuden kannalta sen käyttö on välttämätöntä. Tässä mielessä on tärkeä ymmärtää, että tilastollinen merkittävyys on eri asia kuin kliininen merkittävyys. Klassisena esimerkkinä pidetään verenpaine-mittauksia: systolisen verenpaineen lasku interventoryhmässä kahdella elohopeamillimetrillä voi olla tilastollisesti merkitsevä, mutta yksittäisen henkilön kohdalla muutos ei välttämättä ole merkittävä. Tutkimukseen osallistuvien potilaiden määrällä on suuri merkitys tilastollisen merkittävyyden tutkimisessa. Jos potilasjoukko on pieni, ero kahden

eri ryhmän välillä tulee olla suuri, jotta se olisi tilastollisesti merkitsevä. Kääntäen, jos potilasjoukko on suuri, pienikin ero voi olla tilastollisesti merkitsevä, vaikka ei olisikaan kliinisesti merkittävä.

Yhtä suuri ongelma kuin se, ettei lukija ymmärrä tilastotieteellisiä menetelmiä, on se, etteivät myöskään tutkijat aina hallitse käyttämiään menetelmiä. Väärin käytetty tai valittu menetelmä

myös oleellista, että tutkimusta rakennettaessa on määritetty, voiko ryhmiä ylipäätään verrata keskenään.

Artikkelin tilastollisen analyysin arviointia varten tulee tutkimusjoukon, sen koon ja valikoitumisen olla selkeästi kuvailtu. Myös sen, mitä ja millä mitta-asteikolla tutkimuksessa mitataan, tulisi esittää artikkelissa. Nämä kaikki tekijät vaikuttavat siihen, millaisia tilastomenetelmiä on mielekästä ja mahdollista käyttää aineiston analyysissa. On tärkeää, että artikkelissa on osattu valita aineistoon hyvin soveltuvat tilastomenetelmät. Tilastotieteessä yksinkertainen voi olla kaunista.

Tilastollinen päättely perustuu kahteen periaatteeseen: perusjoukon (populaation) estimointiin ja hypoteesien eli perusolettamusten testaamiseen. Estimoinnissa pyritään löytämään perusjoukon parametreille (esimerkiksi keskiarvolle tai mediaanille) arvio ja määrittää sen luotettavuus. Tällöin tullaan luottamusväleihin. Hypoteesien testauksessa pyritään puolestaan testaamaan hypoteesin paikkansapitävyys, toteamalla esimerkiksi, että kahden potilasjoukon mitatut ominaisuudet eivät eroa toisistaan.

Luottamusväli. Tilastoaineistosta voidaan mitatuille parametreille (esimerkiksi systolinen verenpaine) laskea luottamusväli (confidence interval), jota käyttäen päätelmä tehdään. Lääketieteessä on tavanomaista

TARKISTUSLISTA, 3. osa: Tilastotiede

Menetelmissä on kuvattu tilastolliset menetelmät.	✓
Tunnistan menetelmissä mainitut tilastolliset menetelmät	✓
Aineisto on mielestäni normaalisti jakautunut	✓
Valittu testi on: - parametrisen - ei-parametrisen	✓

voi häiritä tutkimusten tuloksia. Tavallisesti suurimmilla lääketieteellisillä lehdillä on käytettävissä tilastotieteilijä, joka arvioi erikseen tilastollisten menetelmien käyttöä, mutta tästäkin huolimatta virheitä esiintyy. Näiden virheiden merkitys korostuu tilanteissa, joissa tutkimuksen tulos ei pidä sisällään tutkimuskysymyksen kannalta suuria eroja esimerkiksi eri potilasryhmien välillä. On

käyttää 95 %:n luottamusväliä. Sillä tarkoitetaan sitä, että 95 % lasketuista luottamusväleistä peittää perusjoukon tuntemattoman parametrin (esimerkiksi keskiarvon) todellisen arvon. Luottamusvälin sisälle jäänyt vertailuarvo kuvastaa tällöin todellista eroa tutkimusväestössä 95 %:n todennäköisyydellä. Luottamusvälin suuruuteen vaikuttaa paitsi luottamustaso, myös tarkasteltava potilaskoko: mitä suurempi potilasaineisto on, sitä kapeampi on luottamusväli. Toisaalta luottamusvälin suurentaminen esimerkiksi 95 %:sta 99 %:iin leventää luottamusväliä. Luottamusvälin määrittäminen on estimoinnin yhteydessä aina välttämätöntä, jotta tulosten arviointi olisi luotettavaa.

Testin valinta. Tilastolliset testit voidaan karkeasti jakaa parametrisiin ja ei-parametrisiin testeihin. Parametriset testit, kuten Studentin parivertailuun käytetty t-testi, edellyttävät, että mitattava arvo (kuten verenpaine) noudattaa ainakin likimain normaalijakaumaa eli Gaussin kellokäyrää. Tämä tulisi aina tarkistaa ennen ryhmien vertaamista. Ei-parametrisissa testeissä, kuten Mann-Whitneyn U-testi tai Wilcoxonin testi, ei edellytetä, että aineisto on normaalisti jakautunut. On tärkeää ymmärtää periaatteellisesti näiden ero: parametrinen testien käyttö edellyttää aineistolta enemmän oletuksia kuin ei-parametrisen, mutta parametriset testit ovat usein voimakkaam-

pia kuin ei-parametriset, koska tilastolliset merkitsevyydet ovat parametrisillä testeillä helpommin osoitettavissa. Käytännön tasolla ongelma on se, voiko ryhmiä ylipäätään verrata. Vertailua tulkittaessa on huomioitava myös mittarin luotettavuus: esimerkiksi sytokiinin määrityksiin liittynee vähemmän virhelähteitä kuin ahdistuksen määrää kuvaavaan mittariin.

Tilastollinen merkitsevyys.

Kun tutkija on aineiston tarkastelun perusteella päätenyt käyttämään valittua tilastollista testiä, hän usein haluaa selvittää ovatko mitatut tulokset kahdessa eri ryhmässä tilastollisesti merkitsevästi erisuuria. Usein tilastollista eroa kuvataan p-arvolla. Useimmissa lääketieteellisissä tutkimuksissa p-arvoa, joka on pienempi kuin 0,05, pidetään tilastollisesti merkitseväenä. p-arvo ilmoittaa erehtymisriskin. Käytännössä tämä tarkoittaa sitä, että jos p-arvo on pienempi kuin 0,05, väärä päätelmä johtuu pelkästä sattumasta harvemmin kuin kerran 20 tapauksessa tilanteissa, joissa sama tutkimus toteutetaan samoissa olosuhteissa. Liikaa ei kuitenkaan voi korostaa sitä, että tilastollisesti merkitsevä ero ei ole sama kuin käytännön kannalta merkittävä ero.

Tilastotiede on oma tieteenalansa, eikä sen täydellistä hallitsemista tarvitse edellyttää itseltään osatakseen arvioida artikkelin ansioita. Jokaista lääkäriä hyödyttää tutustua muutamaan

yleisimmin käytettyyn tilastolliseen menetelmään. Yleisimmillä menetelmillä katetaan suurin osa lääketieteellisessä tutkimuksessa käytetyistä menetelmistä, jolloin tässä artikkelissa esitettyjen periaatteiden tunteminen on jo suureksi avuksi. ■

OTTO HELVE, LT, lastentautien erikoislääkäri

Helsingin yliopisto
ja HYKS, lastenkliniikka

JORMA KOMULAINEN, LT, dosentti, päätoimittaja

Suomalainen Lääkäriseura Duodecim,
Käypä hoito -yksikkö

JUSSI MERENMIES, dosentti, erikoislääkäri

Helsingin yliopisto
ja HYKS, lastenkliniikka

Eeva PYÖRÄLÄ, dosentti, pedagoginen yliopistonlehtori

Helsingin yliopisto, lääketieteen
koulutuksen tuki- ja kehittämissyksikkö

JUHA HOLOPAINEN, dosentti, silmätautien erikoislääkäri, kliininen opettaja

HYKS, silmäkliniikka