



Hypoteesin testauksesta estimointiin

Estimointi on hypoteesin testausta intuitiivisempi tapa tehdä tilastollisia päätelmiä.

Lääketieteellisessä tutkimuksessa on perinteisesti nojattu hypoteesin testaukseen tilastollisessa päätelyssä.

Yleisin esimerkki hypoteesin testauksesta on ns. nollahypoteesin merkitsevyyden testaus. Siinä on tyypillisesti asetettu ensin nollahypoteesi, kuten kahden ryhmän keskiarvojen oletettu yhtäsuuruus ($H_0: \mu A = \mu B$) tai oletus, että kahden jatkuvan muuttujan välillä ei ole korrelaatiota ($H_0: \rho = 0$). Tämän jälkeen on asetettu vaihtoehdon hypoteesi, joka yleisimmin olettaa, että ryhmien keskiarvot eivät ole yhtä suuret ($H_1: \mu A \neq \mu B$) tai että kahden muuttujan välillä on olemassa korrelaatio ($H_1: \rho \neq 0$). Tämän jälkeen lasketaan havaitun aineiston perusteella p-arvo. Jos se on pienempi kuin ennalta sovittu raja-arvo (yleensä 0,05), hylätään nollahypoteesi ja oletetaan H_1 :n olevan totta. Tällöin päätellään usein, että hoidolla on vaikutusta. Jos p-arvo on suurempi kuin ennalta sovittu raja-arvo, todetaan, että nollahypoteesia ei voida hylätä.

Tilastollinen päätely hypoteesin testauksessa on aina kaksijakoista. Joko vaikutus todetaan eli sen oletetaan olevan olemassa tai sitten oletettua vaikutusta ei voida osoittaa. Tämä on kuitenkin epäintuitiivinen tapa tehdä tilastollisia päätelmiä. Biolääketieteessä kaikki vaikutukset ovat aina jatkumoitteita. Mikään vaikutus ei ole koskaan täysin puuttuva tai kaksijakoinen.

Käytännöllisempi näkemys on ajatella, että kaikki vaikuttaa kaikkeen ja kyse on vain siitä, minkä suuruisesta vaikutuksesta on kyse. Tietenkin aitoja nolla-vaikutuksia voidaan ajatella olevan, ku-

ten suolaveden vaikutus haimasyöpään, mutta tämäntyyppisiä tutkimuskysymyksiä harvemmin laaditaan. Lääketiede perustuu viime kädessä hyvin monimutkaisiin biokemiallisiin ja fysiologisiin tapahtumiin, ja siksi todelliset nollavaikutukset lienevät hyvin harvinaisia.

Estimointi on hypoteesin testausta intuitiivisempi tapa tehdä tilastollisia päätelmiä biolääketieteessä. Tutkittaessa hoitovaikutuksia ja riippuvuus-suhteita on käytännöllisempää ajatella, että jonkinlainen vaikutus on aina olemassa. Kyse on vain siitä, minkä suuruinen ja millä suureilla se on. Estimointi tarkoittaa parhaan arviointia etsimistä jollekin muuttujalle, kuten ryhmien keskiarvojen väliselle erotukselle (1).

Viime kädessä estimointi nojaa luottamusvälien tulkintaan. Kyseessä ei siis suinkaan ole mikään uusi tai erilainen toimintatapa, koska p-arvo ja luottamusvälit ovat vain saman kolikon kaksi eri puolta. Kyse on ennen kaikkea tulkintatavasta, miten luottamusvälejä hyödynnetään. Luottamusvälit ovat keskeinen osa tulostuottajien raportointia, mutta jos luottamusvälin tulkinta perustuu arvioon siitä, sisältyykö arvo 0 tai 1 siihen, kyseessä on kuitenkin hypoteesin testaus eikä luottamusvälejä ole hyödynnetty oikeapoisesti.

Estimoinnissa tarkoitetaan tulkittua luottamusvälejä joukkona sellaisia arvoja, joiden kanssa aineisto on yhtenevä.

Keskeistä on tulkita näitä arvoja myös suhteessa kliinisesti tai potilaan kannalta merkityksellisiin arvoihin. p-arvoja tarvitse raportoida periaatteessa ollenkaan.

Estimoinnin hyvä puoli on se, että silloin vältetään virheelliset tulokset, kuten ”ryhmien välillä ei ollut eroa”. Tätä ei voida koskaan todeta perinteisen frekventistisen tilastotieteen keinoin.

Jos tutkimuksessa ei havaita tilastollisesti merkitsevää löydöstä, hypoteesin testauksessa on perinteisesti todettu, että ryhmien välillä ei ollut eroa tai tilastollisesti merkitsevää eroa ei todettu. Jos tulosta tulkitaan estimoinnin näkökulmasta, voidaan todeta esimerkiksi lääkkeen tehosta, että

aineiston perusteella voidaan sulkea pois 0,12 mmol/l suurempi lasiku kolesteroliarvossa eli teho lienee vaatimaton.

Lääketiede perustuu viime kädessä hyvin monimutkaisiin biokemiallisiin ja fysiologisiin tapahtumiin, ja siksi todelliset nollavaikutukset lienevät hyvin harvinaisia. Tämän vuoksi estimointi on intuitiivisempi tapa tehdä päätelyä verrattuna hypoteesin testaukseen. ●

ALEKSI REITO
aleksi@reito.fi

KIRJALLISUUTTA

- 1 Claridge-Chang A, Assam PN. Estimation statistics should replace significance testing. *Nat Methods* 2016;13(2):108–9.