

Näyttöön perustuvasta lääketieteestä – Lääketieteellisen tiedon kriittistä tarkastelua

Aleksi Pajunen

Lääketieteen kandidaatti

Clinicum

Helsinki 13.12.2016

Tutkielma

aleksi.pajunen@helsinki.fi

Ohjaaja: Pekka Louhiala

HELSINGIN YLIOPISTO

Lääketieteellinen tiedekunta

HELSINGIN YLIOPISTO – HELSINGFORS UNIVERSITET

Tiedekunta/Osasto – Fakultet/Sektion – Faculty Lääketieteellinen tiedekunta		Laitos – Institution – Department Clinicum	
Tekijä – Författare – Author Aleksi Pajunen			
Työn nimi – Arbetets titel – Title Näyttöön perustuvasta lääketieteestä – Lääketieteellisen tiedon kriittistä tarkastelua			
Oppiaine – Läroämne – Subject Lääketieteen etiikka			
Työn laji – Arbetets art – Level Syventävä tutkielma		Aika – Datum – Month and year 13.12.2016	Sivumäärä - Sidoantal - Number of pages 42 + 1
Tiivistelmä – Referat – Abstract <p>Näyttöön perustuva lääketiede käsitteenä esiteltiin suurelle yleisölle ensimmäisen kerran vuonna 1992, jonka jälkeen se on saanut laajaa kannatusta ja herättänyt runsasta keskustelua.</p> <p>Tässä kriittisessä kirjallisuuskatsauksessa tarkastelen EBM:n määrittelyä sekä sille esitettyjä argumentteja puolesta ja vastaan. Teen huomion, että käytännön tasolla EBM tarkoittaa usein tutkimusmetodologioiden arvottamista, jossa RCT-tutkimusten tuottama tieto nähdään arvokkaampana kuin muiden tutkimusten. Tarkastelen RCT:in puolesta esitettyjä perusteluja ja esitän, että satunnaistamisen arvoa tutkimuskäytössä on ylikorostettu. Lisäksi RCT:eihin liittyy omat virhelähteensä ja riskinsä, kuten esimerkiksi tekninen harha, joka saattaa suosia lääkkeellisiä interventioita.</p> <p>Evidenssihierarkian sijaan tutkimuksia ja laadullisesti erilaisia lääketieteellisen tiedon muotoja tulisikin tarkastella nykyistä kokonaisvaltaisemmin.</p> <p>(96 sanaa)</p>			
Avainsanat – Nyckelord – Keywords Evidence-Based Medicine, Randomized Controlled Trial, Philosophy of Medicine			
Säilytyspaikka – Förvaringställe – Where deposited			
Muita tietoja – Övriga uppgifter – Additional information			

1 Johdanto	1
2 Käsitteen määrittely	3
3 Filosofinen katsaus	9
3.1 Mitä ilmaus ”näyttöön perustuva” tarkoittaa?	9
3.2 Mitä on lääketiede?	13
4 Näyttöön perustuva lääketiede – hyviä ja huonoja puolia.....	15
4.1 Esimerkki EBM:n eduista: rytmihäiriölääkkeet sydämen vajaatoiminnassa	15
4.2 Tutkimuksen virhelähteiden huomioiminen koasetelmaa laatiessa	16
4.3 Satunnaistettu kontrollitutkimus – laadukkainta tietoa?	20
4.4 Retroaktiivisen rukouksen vaikutus – taustateorian ja tutkimuksen välisestä yhteydestä	27
4.5 Juridinen tapauselostus: PSA-seulonta ja hoitolinjat	29
5 Psykologinen näkökulma: Vividness-hypoteesi	31
6 Mitä EBM:n tilalle?	33
6.1 Hierarkiaa laajempaa tutkimusten arviointia	33
6.2 Kasuistinen lääketiede	34
7 Yhteenveto.....	37
8 Johtopäätökset.....	38
9 Lähdeluettelo	40

1 Johdanto

Näyttöön perustuva lääketiede (evidence-based medicine, EBM) on ollut runsaan keskustelun aiheena jo 1990-luvulta alkaen. Lääketieteellisissä tiedekunnissa opetussuunnitelmaan kuuluu usein EBM-kurssi, jolla opetellaan tieteellisten julkaisujen arviointia ja useilla kliinisillä kursseilla puhutaan ”EBM:n mukaisesta hoidosta”. Tätä taustaa vasten tuntuukin yllättävältä, että oman kuuden vuoden koulutusohjelmani aikana en ole kuullut kerrottavan mitä tämä näyttöön perustuva lääketiede tarkalleen on ja erityisesti mitä on sen vastakohta, näyttöön perustumaton lääketiede.

EBM-käsitteellä tuntuukin olevan nykypäivän keskustelussa eräänlainen ”mahtisanan” rooli: jokin hoito tai valinta voidaan perustella tai kyseenalaistaa automaattisesti sanomalla sen olevan EBM:n mukaista tai vastaista ilman että kerrotaan, mitä tämä EBM edes on eli mitä kriteereitä hoidon olisi tullut täyttää tai miksi valitut kriteerit eivät kelpaa.

Tämä tutkielma on tutkimusmenetelmältään kriittinen kirjallisuuskatsaus, jossa tarkastelen EBM-käsitteen syntyä ja siihen liittyvää tieteellistä keskustelua. Tutkielman kysymyksenasetteluna on: miten käsite EBM määritellään, tarkoittaako se käytännön lääketieteessä samaa kuin määritelmänsä ja onko näille näkökulmille kestäviä perusteita? Katsauksen aineistona ovat keskeiset aihepiirin artikkelit.

Päädyn kyseenalaistamaan käsitteen käytön mielekkyyden ja teen johtopäätöksen, että käytännön tasolla EBM:llä tyypillisesti tarkoitetaan tietynlaisen lääketieteellisen tiedon arvottamista muita ylemmäksi. Kyseenalaistan myös tämän hierarkkisen ajattelun mielekkyyden.

Käyn myös läpi satunnaistetun tutkimusasetelman hyviä ja huonoja puolia sekä teorian ja evidenssin välistä suhdetta. Lopuksi esittelen, millaisia malleja on tarjottu EBM:ksi kutsutun ajattelun hyvien puolien liittämiseksi osaksi lääketieteen tutkimusta ja lääkärin profession harjoittamista.

Tavoitteenani on ollut lähestyä aihetta ennen kaikkea kyselevällä mielenlaadulla. Vaikka tämä tutkielma useissa kohdin ottaa hyvin kriittisen lähestymistavan tarkastelemaansa aiheeseen, se johtuu halusta esittää kysymyksiä juuri niistä alueista, joissa EBM-ajattelulla on vielä kehitettävää. Näen, että EBM-ajatuksilla on paljon annettavaa lääketieteeseen, mutta niiden käyttäminen ainoana toimintamallina niitä kyseenalaistamatta tai kehittämättä voi helposti olla yhtä vaarallista kuin niiden kategorinen hylkääminenkin.

Kaikki sitaatit ovat itse kääntämiäni.

2 Käsitteen määrittely

Käsitteen ”näyttöön perustuva lääketiede” syntyperää avaa artikkeli ”Everything you wanted to know about Evidence-based Medicine”. Käsitteen on alun perin kehittänyt McMaster yliopiston Gordon Guyatt, joka halusi koulutusohjelman antavan työkaluja lääketieteellisten tutkimusten arviointiin ja soveltamiseen. Hänen alkuperäinen ehdotuksensa ”tieteellinen lääketiede” herätti voimakasta vastustusta. Perustutkimuksen tekijät ymmärrettävästi kritisoivat vihausta, että heidän harjoittamansa lääketiede olisi epätieteellistä. Toinen ehdotus, ”näyttöön perustuva lääketiede”, sen sijaan löi itsensä menestyksekkäästi läpi.¹ Tällä kertaa kukaan ei kysynyt harjoittavatko perustutkimuksen tekijät näyttöön perustumatonta lääketiedettä.

Suurelle yleisölle käsite ”evidence-based medicine” esiteltiin vuonna 1992 artikkelissa ”Evidence-Based Medicine – A New Approach to Teaching the Practice of medicine”.²

Artikkelin johdannossa käsite määriteltiin seuraavasti: ”Lääketieteen harjoittamiseen on muodostumassa uusi paradigma. Näyttöön perustuva lääketiede vähentää intuition, ei-systemaattisen kliinisen kokemuksen ja patofysiologisten perustelujen roolia kliinisessä päätöksenteossa ja painottaa kliinisistä tutkimuksista saadun näytön merkitystä.”²

Perusta paradigman muutokselle oli kirjoittajien mukaan tieteellisen tutkimuksen roolin muuttuminen lääketieteen maailmassa edeltävän 30 vuoden aikana: vuonna 1960 satunnaistettu kliininen koe oli harvinaisuus, kun taas vuonna 1992 se nähtiin välttämättömyytenä lääkkeen markkinoille tuomiseksi. Kirjoittajien mukaan satunnaistettuja kokeita laaditaan jatkuvasti enemmän ja niitä sovelletaan sellaisiin lääketieteen osa-alueisiin, joissa niitä ei aiemmin ole käytetty (mm. kirurgiaan, diagnostisiin kokeisiin ja ennusteen laatimiseen).²

Kirjoittajat esittävät vanhan lääketieteen paradigman tausta-oletuksiksi seuraavat:

1. Kliinisen kokemuksen tuomat ei-systemaattiset havainnot ovat hyväksyttävä tapa rakentaa ja ylläpitää lääkärin osaamista potilaan ennusteesta, diagnostisten testien laadusta ja hoitojen vaikuttavuudesta.
2. Ymmärrys patofysiologiasta ja sairauksien perustason mekanismeista ovat riittävä perusta (sufficient guide) kliinisen työn harjoittamiseen.
3. Kattava lääketieteellinen koulutus ja maalaisjärki ovat riittäviä edellytyksiä uusien testien ja hoitojen arviointiin.
4. [Yksilön] lääketieteen asiaosaaminen ja kliininen kokemus ovat riittävä pohja hoitosuosituksen laatimiseksi.

Esitellyn perinteisen paradigman kuvataan painottavan [lääke]tieteellisiä auktoriteetteja. Tätä mallia on myöhemmin kritisoitu epärealistiseksi kuvaksi ”perinteisestä lääketieteestä”.²

Uudelle paradigmalle esitetään seuraavat taustaoletukset:

1. Kliininen kokemus ja kliinisen vaiston kehittäminen ovat välttämättömiä hyväksi lääkäriksi kehittymiselle, koska kaikkia lääketieteen osa-alueita ei voida testata asianmukaisesti. Kuitenkin datan kerääminen toistettavalla ja vääristämättömällä tavalla parantaa luottamusta potilaan ennusteeseen ja tilanteeseen kohdistuvan tiedon osalta.
2. Patofysiologisten mekanismien ymmärtäminen on tärkeää, mutta sitä ei tule käyttää päätöksenteon tukena. Patofysiologisten periaatteiden varaan rakennettu päättely voi johtaa epätäydelliseen diagnoosiin ja/tai hoitoon, koska patofysiologinen päättely voi olla väärässä.
3. Lääkärit, joiden antama hoito perustuu hoidon taustalla olevan näytön (evidence) ymmärtämiseen, antavat potilaille laadukkaampaa hoitoa.

Vaikka ”asianmukaista testaamista” ei taustaoletuksissa määritellä tarkemmin, voidaan myöhemmin esitetyn perusteella päätellä tämän tarkoittavan satunnaistettua kontrollitutkimusta (Randomized controlled trial, RCT). Samoin kohdassa 2. eksplisiittisesti ainoastaan kritisoidaan patofysiologista päättelyä, mutta muun

artikkelissa esitetyn perusteella tämä ilmeisimmin johtuu halusta painottaa (mitattavissa olevien) päätetapahtumien merkitystä hoidon valintakriteerinä.²

Uuden paradigman taustaoletuksia ei tässä kohtaa tarkastella kriittisesti. Kirjoittajat eivät myöskään esitä empiiristä dataa uuden paradigman tueksi. He eivät huomioi mahdollisuutta, että systemaattisesti kerätty näyttö voisi myös olla väärässä vaan sen luotettavuus nähdään itsestään selvänä. Eräs toistuvista kritiikeistä EBM:ä kohtaan kohdistuu 3. oletukseen, joka on otettu aksiomaattiseksi taustaoletukseksi tarjoamatta minkäänlaista empiiristä dataa oletuksen todenperäisyydestä³.

Artikkeli esittää myös ehdotuksia kriteereistä, joilla valitaan metodologisesti laadukkaita tutkimuksia:

- Diagnoosi – Onko diagnostinen testi arvioitu potilasryhmällä, joka sisältää riittävän laajasti sekä lieviä että vakavia oireita omaavia potilaita, sekä hoidetun että hoitamattoman sairauden kanssa? Onko lisäksi otettu mukaan yksilöitä, joilla on tutkittavaan diagnoosiin helposti sekoittuva eri tauti? Onko mukana erillinen sokkoutettu vertailu kyseisen diagnostiikan ”kultaiseen standardiin”?
- Hoito – Onko potilaiden hoito satunnaistettu? Onko lopputuloksessa huomioitu kaikki kokeeseen otetut potilaat?
- Katsausartikkelit – Onko kriteerit, joilla artikkelit valikoitiin tutkimukseen, esitelty selvästi?

Kirjoittajat esittävät, että myös monimutkaisempia kriteereitä voidaan käyttää EBM-osaamisen lisääntyessä. Kriteerit tulisi valikoida ”nihilismiä painottaen (jos koe ei ole satunnaistettu, se on turha eikä tuota käyttökelpoista tietoa.)”²

Artikkeli myös kertoo, että EBM:ä on kritisoitu siitä, että se hylkää kliinisen kokemuksen ja intuition. Kirjoittajat kuitenkin toteavat, että kliininen tieto ja intuitio voi antaa dataa, joka myöhemmin todennetaan asianmukaisella testillä. Näytön puuttumisen vuoksi taas voi olla välttämätöntä nojautua patofysiologiseen päättelyyn kunnes parempaa näyttöä on tarjolla.² Kirjoittajien ajattelumallissa kliininen tieto ja patofysiologiset mekanismit

ovat siis alempiarvoisia kuin heidän hierarkiassaan korkeammalle sijoittuvat tutkimusmenetelmät.

Kirjoittajat kertovat myös, että EBM:n eduista lyhyellä aikavälillä ei ole näyttöä. Sen sijaan lyhyen aikavälin tutkimukset ovat osoittaneet, että ”EBM-taitoja” voidaan opettaa opiskelijoille sekä lääkäreille ja että EBM-koulutetut opiskelijat ovat lääkäreinä paremmin ajan tasalla uusista hoitokäytännöistä. Kirjoittajat itsekin myöntävät, että EBM:n puolustaminen ilman näyttöä sen toimivuudesta voi vaikuttaa sisäiseltä ristiriidalta. Tähän vastauksena esitetään, että ”definiitiivisen näytön” puuttuessa tulee turvautua ”heikompaan näyttöön” kuten biologiseen päättelyyn. EBM:n puolesta oletetaan aksiomaattisesti, että ajan tasalla hoitokäytännöistä olevat lääkärit tuottavat parempia hoitotuloksia.² Kuitenkaan tässä yhteydessä ei tarkastella komplementaarista aksiomaa - että ajan tasalla hoitokäytännöistä olevat lääkärit eivät tuota merkityksellisesti parempia hoitotuloksia – tai sitä, millä perusteella he ovat päätyneet ensimmäiseen näistä kahdesta.

Artikkelissa siis esitetään väite, että lääketieteellisten päätösten tulisi perustua laadukkaaseen tutkimustietoon. Lisäksi esitetään väite, että vain satunnaistetuista tutkimuksista saadaan laadukasta tietoa.

Tarkastelen tarkemmin EBM:n hyviä ja huonoja puolia myöhemmin tässä tutkielmassa, mutta kuvaan ensin alkuperäiselle JAMA:in artikkelille esitetyn kritiikin pohjalta vuonna 1996 julkaistua v EBM:n käsitettä tarkentava ja uudelleenmäärittelevää kirjoitusta ”Evidence based medicine: what it is and what it isn't”.⁴

Kirjoittajat määrittelevät artikkelinsa alussa EBM:n seuraavasti: ”Näyttöön perustuva lääketiede on tietoista, eksplisiittistä ja arvostelukykystä parhaan näytön käyttöä yksittäisten potilaiden hoitoa koskevissa päätöksissä.” EBM:n harjoittamisen kuvataan olevan yksilön kliinisten taitojen yhdistämistä parhaaseen systemaattiseen tutkimustietoon. Tutkimustiedossa paikkansa on myös lääketieteen perustieteillä, mutta kirjoittajat sanovat keskeisimmän painoarvon olevan ”potilaskeskeisessä

kliinisessä tutkimuksessa”. Näitä ovat diagnostisten testien tarkkuus, ennustemarkkereiden voimakkuus sekä hoitomuotojen tehokkuus ja turvallisuus.⁴

Artikkelin mukaan hyvä lääkäri käyttää sekä kliinistä kokemustaan että ulkoista tutkimusnäyttöä eikä kumpikaan yksinään ole riittävää. Artikkelin kuvaus, että ilman kliinistä asiantuntemusta lääketiede joutuu ”evidenssin tyrannisoimaksi”, koska tutkimusnäytön soveltuvuutta yksittäiselle potilaalle ei arvioida kunnolla. Jos taas uusin tutkimusnäyttö jää huomioimatta, on vaarana hoidon käyminen vanhanaikaiseksi.⁴

Lisäksi artikkelissa vastataan EBM:ä kohtaan esitettyyn kritiikkiin tarkentamalla, mitä EBM ei ole. Kaksi keskeistä väitettä ovat, että EBM:ä harjoitetaan tai että sen toteuttaminen käytännössä on mahdotonta. Kirjoittajat viittaavat tutkimuksiin, joissa on osoitettu merkittäviä lääkärikohtaisia eroja hoitokäytänteissä, minkä he tulkitsevat osoittavan, että EBM:ä ei koko lääkärikunnassa harjoiteta. Jälkimmäiseen kritiikkiin ”EBM:ä voidaan harjoittaa norsunluutorneista käsin, ei kentällä” vastataan antamalla esimerkkejä tiimeistä, jotka ovat kohdennetulla tiedonhaulla onnistuneet EBM:ssä.⁴

Kirjoittajat varoittavat myös käyttämästä EBM:ä ”keittokirjalääketieteen” työkaluna, koska tällöin potilaan yksilöllinen tilanne jää huomioimatta. Artikkelissa myös todetaan, että EBM:n ottaminen poliittisen päätöksenteon välineeksi on sen väärinkäyttöä. Selkeytyksenä aiempaan, sanotaan myös selkeästi, että näyttöön perustuvaan lääketieteeseen kuuluu monenlaisia tutkimusasetelmia RCT:in lisäksi. RCT:in arvo on erityisesti hoidon valikoinnissa, koska näissä subjektiivinen näyttö voi johtaa vääriin positiivisiin tuloksiin tai muuten viallisiin tulkintoihin. Koska RCT ”ei harhaan johda tutkijoita”, on se paras kultainen standardi tällaisten asioiden tutkimisiin.⁴

Näiden kahden artikkelin antamien määritelmien lisäksi keskeinen osa EBM-keskustelua ovat käytännössä olleet evidenssin hierarkiat. Näitä on esitelty useampia, mutta yhteistä niille on erilaisten tutkimusmetodien – ja siten niistä saatavan tiedon – luokittelu, jossa tulisi aina pyrkiä hierarkian mukaan laadukkaimman mahdollisen tiedon tuottamiseen ja seuraamiseen. Näissä hierarkioissa RCT nähdään ”kultaisena standardina” ja

yksittäisen klinikon yksittäiset tapauselostukset taas huonolaatuisimpana, ja vain anekdotaalisena evidenssinä.

EBM:n puolestapuhujat ovat argumenteissaan painottaneet tämän hierarkian etuja erityisesti subjektiivisuuden eliminoimiseksi lääketieteellisessä tutkimuksessa ja tarjonneet esimerkkejä tapauksista, joissa nimenomaan RCT on tuottanut keskeistä hoidon laatua parantavaa dataa. EBM:n kritiikki on karkeasti jaettavissa kritiikiksi evidenssihierarkiaa kohtaan, termin EBM merkitykseen, käyttöön ja tarpeellisuuteen sekä haasteisiin EBM:n ottamisessa käytäntöön esitetystä muodostaan. Tarkastelen näitä kaikkia tarkemmin myöhemmissä kappaleissa.

3 Filosofinen katsaus

3.1 Mitä ilmaus ”näyttöön perustuva” tarkoittaa?

Olen nyt tarkastellut sitä, miten EBM on alun perin määritelty. Seuraavaksi tarkastelen, mitä EBM:n muodostavat kolme sanaa tarkoittavat ja kuinka ne kytkeytyvät toisiinsa.

Artikkelissaan ”The concept of evidence in evidence-based practice” Tone Kvernbekk kirjoittaa filosofisia näkökulmia evidence based practice –keskusteluun kasvatustieteissä.⁵ Hänen esittelemänsä Achinsteinin filosofinen evidenssin käsite⁶ on hyvin sovellettavissa myös lääketieteeseen.

Achinstein kuvaa evidenssin yksinkertaisimman määritelmän olevan ”se, mikä oikeuttaa uskomuksen” eli jotain, mikä tukee hypoteesin todenperäisyyttä tai väärässä olemista. Evidenssin rooli on teorian tukeminen (support).⁶ Tähän tukemiseen kykenevät niin faktat, kokemukset kuin data useissa eri muodoissaan ja ne soveltuvat kaikki evidenssiksi, jos ne täyttävät tämän tukemisen tarkoituksensa luotettavasti. Kvernbekk toteaaakin, että evidenssin konseptissa ’itsessään’ ei ole mitään, mikä nostaisi RCT:in tai kvantitatiivisen datan arvokkaammaksi kuin muun datan (lainausmerkit omat).⁵

Evidenssi on siis vain tukea hypoteesille. Kvernbekk esittelee hypoteettis-deduktiivisen näkemyksen: Hypoteesista H on voitava johtaa (deduce) ennuste evidenssin E olemassaololle ja/tai havaitsemiselle. Tämä evidenssin ja hypoteesin välinen suhde on välttämätön, muttei riittävä ehto, jotta E voi olla evidenssi H:lle. Tästä kuitenkin seuraa väistämättä, että E toimii evidenssinä kaikille, jopa keskenään ristiriitaisille hypoteeseille, joista voidaan johtaa ennuste E:n havaitsemiselle.⁵

Koska lääketieteessä potilaan hyvinvoinnin ja terveyden lisääminen on keskeinen päämäärä, kohdistuu ymmärrettävästi suurin mielenkiinto evidenssiin, joka osoittaa jonkin intervention syy-yhteyden potilaan tilan paranemiseen, siis kausaaliteettia tukevaan evidenssiin. Kirstin Borgerson käsittelee kausaaliteetin todistamista lääketieteessä. Hänen mukaansa lääketieteelliset kausaatiot tyypillisesti jakautuvat

kahteen luokkaan, mekanistisiin ja probabilistisiin. Mekanistiset kausaatiot nousevat biotieteistä, genetiikasta, fysiologiasta ja muista lääketieteen perustieteistä. Probabilistiset kausaatiot taas kuvaavat kahden tekijän välisen assosiaation vahvuutta ja nousevat epidemiologisista tutkimusasetelmista.⁷

Borgerson myös esittelee Bradford-Hillin yhdeksän kriteeriä kausaatiolle: assosiaation vahvuus, ajallinen yhteys, johdonmukaisuus (consistency), teoreettinen uskottavuus (plausibility), koheesio, spesifisyys, annos-vaste –suhde, kokeellinen evidenssi ja yhdenmukaisuus/vastaavuus (analogy).⁷

Näistä mekanistisiin kausaatioihin kytkeytyvät vahvasti ajallinen yhteys, teoreettinen uskottavuus, koheesio ja kokeellinen evidenssi. Probabilistisiin kausaatioihin taas kytkeytyvät vahvasti assosiaation vahvuus, johdonmukaisuus, spesifisyys ja annos-vaste –suhde. Borgersonin mukaan on epäselvää, millä kriteerillä EBM-ajattelussa on päätetty nostaa probabilismiin kytkeytyvät kausaation kriteerit yli mekanistisiin kytkeytyvien, vaikka kaikki niistä ovat päteviä evidenssin luokkia kausaatiolle.⁷ Kysymys on erinomainen ja Achinsteinin evidenssinkäsityksen ”se, mikä oikeuttaa uskomuksen” valossa tällainen jaottelu ei näytä oikeutetulta.

Evidenssi itsessään voi siis olla hyvin monimuotoista. Sana ”perustua” puolestaan kuvaa evidenssin ja lääketieteen välistä kytköstä. Tarkalleen ottaen lääketiede perustuu hypoteeseille, jotka perustuvat evidenssille. Toisin sanoen evidenssin tiedolliselle sisällölle on aina tehty jonkinlainen tulkinta ennen sen soveltamista käytäntöön. Kvernbekk antaa tästä hyvän esimerkin: veritahra vaatteissa ja kaksi todistajaa ovat evidenssiä, mutta teoria on, että hovimestari on syyllinen. On myös tärkeää muistaa, että evidenssi ei ole konklusiivista; voi paljastua, että hovimestari onkin syytön.⁵

Lisäksi Kvernbekk painottaa, että evidenssi usein johtaa teoriaan, mutta suinkaan aina teoria ei ole suoraan johdettavissa evidenssistä. Tästä esimerkkinä Newtonin gravitaatioteoria: data, jota Newtonilla oli saatavilla ei itsestään selvästi näytä osoittavan nimenomaan gravitaatioteoriaan vaan pikemminkin herättää Newtonissa kysymyksen ”mikä näkymätön tekijä selittää tällaiset löydökset”. Toinen tapa tarkastella

tätä ilmiötä on alimääräytyneisyys, jonka mukaan evidenssi johtaa teoriaan aina vain osittain. Teorian muodostumiseen tarvitaan evidenssin lisäksi myös huomattava määrä aiempia teorioita, kokemuksia ja näkemyksiä.⁵ Evidenssin ja hypoteesin välinen suhde ei siis koskaan muodostu tyhjiössä 'de novo'. Vaikka haluaisimme erottaa uuden teoreettisen mallin luomisprosessin tieteen aiemmasta kulusta, ei se käytännössä ole mahdollista ja tämän tosiasian kieltäminen johtaa nopeasti epärealistiseen käsitykseen evidenssin omavoimaisuudesta.

Tämän lisäksi, kuten hypoteettis-deduktiivinen malli osoittaa, sama evidenssi voi tukea lukuisia eri teorioita, minkä vuoksi muun tietämyksen ja kontekstin huomioiminen on itse asiassa välttämätöntä eri hypoteesien arvottamiseksi. Kvernbekk ehdottaakin ilmaisumuodoksi Erwinin ja Siegelin näkemystä: E vahvistaa H:ta, jos se lisää uskottavasti syitä ottaa H totena, eikä samanaikaisesti nosta samassa määrin muita mielekkäitä hypoteeseja⁸.

Medisiinisessä keskustelussa evidenssin ja käytännön työn välistä suhdetta on käsitellyt Mark Tonelli artikkelissaan "The philosophical limits of evidence based medicine". Hän kuvaa, kuinka lääketieteen traditio nojaa potilaan ja lääkärin suhteeseen ja lääkärin tavoitteeseen tarjota senhetkisen tietämyksen valossa parasta hoitoa omalle potilaalleen. EBM:n hän sanoo tutkivan ennen kaikkea keskimääräisiä vaikutuksia väestötasolla, jotka ovat terveysviranomaisten kannalta keskeisiä, mutta jotka eivät tarjoa vastauksia tähän lääkärin tärkeimpään tavoitteeseen.⁹

Toinen Tonellin esiin nostama ongelma liittyy ilmiöiden mitattavuuteen. Yksilön status voi sisältää johtolankoja diagnoosin tekemiseen, jotka kokenut klinikko intuitiivisesti havaitsee, mutta joita ei kyetä pukemaan eksplisiittisiksi diagnostisiksi kriteereiksi. Tai joiden numeerinen arvottaminen tulee luultavasti aina olemaan hataraa tai subjektiivista, kuten kipu. Nämä tekijät selvästi vaikuttavat potilaan tilaan, diagnoosiin ja ennusteeseen. Mutta jos ne eivät luontevasti taivu numeeriseen esitysmuotoon, jää EBM-kelpoisen tiedon ja käytännön lääketieteen välille aukko, jota EBM ei itse asettamiensa rajausten vuoksi kykene kuroma umpeen.⁹

Kenties merkittävin Tonellin esittämä kritiikki on kuitenkin eettinen: moraalifilosofiassa on esitetty olevan laadullinen ero faktojen ja arvojen välillä, jonka vuoksi arvoja ei koskaan voi johtaa suoraan datasta. Tämän esitetään johtuvan faktojen ja arvojen tiedollisesti erilaisesta luonteesta, jonka vuoksi ne eivät ole suoraan kytköksissä toisiinsa.⁹

Esimerkiksi evidenssin tukiessa hypoteesia: ”Jos haluat maksimoida todennäköisyyden tulokselle A, tee B”, ei evidenssi eikä hypoteesi missään kohtaa ota kantaa siihen, onko A:n maksimointi ylipäättään haluttavaa tai tavoittelemisen arvoista. Lauseiden ”A tulee maksimoida” ja ”näin A maksimoidaan” ainoa yhtymäpiste on, että molemmat sattuvat käsittelemään A:ta. Ne ovat kuitenkin toisistaan irrallisia eikä kumpikaan ole johdettavissa toisesta.

Tätä tiedon ja arvojen eroa kutsutaan ”on – pitäisi” –eroksi (is – ought –gap), josta seuraa, että edes datan määrän lisääminen ei kerro meille, mitä tietyssä tilanteessa pitäisi tehdä⁹. Lääketieteessä tämä kysymys nousee esiin esimerkiksi eutanasiakeskustelussa, jos kykenemme sekä pidentämään elämää, että lyhentämään kärsimyksen kestoja, mutta emme voi tehdä molempia yhtä aikaa. Lääkärin tiedollinen kykeneminen kumpaan tahansa vaihtoehtoon ei vielä tarjoa vastausta siihen, kumpi niistä tulisi valita.

Tämän tarkastelun pohjalta voidaan todeta, että evidenssi on jotain, mikä nostaa tietyn hypoteesin uskottavuutta suhteessa muihin hypoteeseihin, mutta evidenssin käsite ei vielä anna perusteita suosia tietynlaista evidenssiä tai tutkimusmetodologiaa. Toisena johtopäätöksenä on, että ”perustua” on erittäin voimakas ilmaus, joka ei kuvaa luontevasti evidenssin translaatiota käytännön lääketieteeseen. Luontevampi sana voisi olla esimerkiksi ”tukea” (support), mutta kuten seuraavassa kappaleessa totean, voidaan perustellusti koko etuliite ”näyttöön perustuva” nähdä tarkoituksettomana lääketieteeseen liitettynä.

3.2 Mitä on lääketiede?

Daniel Callahan on määritellyt lääketieteen seuraavasti: ”Mielestäni lääketiede on humanistinen oppiaine, joka käyttää tiedettä saavuttaakseen tavoitteitaan. Tämän sanomalla haluan tehdä eron yleiseen uskomukseen, että lääketiede on tieteellinen oppiaine, jolla on joitain sekundaarisia humanistisia arvoja ja päämääriä.”¹⁰

Callahanin näkökulma antaa arvon tieteelliselle tutkimukselle, analyyttiselle ymmärrykselle ihmisen ja sairauksien fysiologiasta sekä interventioden mekanistisille ja probabilistisille vaikutuksille. Kuitenkin hän huomioi, että lääketiede ei ole pelkästään ”kylmää dataa” vaan kyseessä on kokonaisvaltainen potilaan terveyden ja hyvinvoinnin lisääminen, jossa tieteellisellä ymmärryksellä on välinearvo tämän tärkeämmän päämäärän tavoittamiseksi. Tällä on samankaltaisuutta ”on – pitäisi” –ajattelun kanssa: lääketieteessä ensisijaisia ovat tavoittelemisen arvoiset päämäärät. Tieto siitä, miten päämäärät saavutetaan, on tärkeää, mutta toissijaista tähän ensisijaisuuteen nähden ja irrallista siitä.

Mutta mitä sitten on ”näyttöön perustuva lääketiede” ja miten se eroaa lääketieteestä? Onko olemassa erillinen ”ei-näyttöön perustuva lääketiede”, jota EBM:n vastustajat harjoittavat?

Suomalainen näkemys lääkäriydestä kiteytyy parhaiten lääkärinvalaan, jossa sanotaan seuraavaa: ”Päämääränäni on terveyden ylläpitäminen ja edistäminen, sairauksien ehkäiseminen sekä sairaiden parantaminen ja heidän kärsimyksensä lievittäminen. Työssäni noudatan lääkärin etiikkaa ja käytän vain lääketieteellisen tutkimustiedon tai kokemuksen hyödyllisiksi osoittamia menetelmiä. Tutkimuksia ja hoitoja suositellessani otan tasapuolisesti huomioon niistä potilaalle koituvan hyödyn ja mahdolliset haitat. Pidän jatkuvasti yllä korkeaa ammattitaitoani ja arvioin työni laatua.”¹¹

Tämä on hyvin linjassa näyttöön perustuvan lääketieteen määritelmän kanssa: ”Näyttöön perustuva lääketiede on tietoista, eksplisiittistä ja arvostelukykyistä parhaan näytön käyttöä yksittäisten potilaiden hoitoa koskevissa päätöksissä.”⁴ Eyal Shahar

kuitenkin kritisoi tätä näyttöön perustuvan lääketieteen määritelmää tekemällä luontevan johtopäätöksen, että ”lääkäri, joka ei harjoita EBM:ä ei siis käytä parasta tarjolla olevaa näyttöä tietoisesti, eksplisiittisesti ja arvostelukykyisesti.”¹² Tällaista käytäntöä ei kuitenkaan lääkärinvalan perusteella voi kutsua ylipäätään lääketieteen harjoittamiseksi.

Nämä sanalliset määritelmät eivät siis erota lääketiedettä näyttöön perustuvasta lääketieteestä. Konkreettisenä erona ”perinteiseen” lääketieteeseen onkin ennen kaikkea tapa suhtautua lääketieteelliseen tietoon ja arvottaa sitä tutkimusmetodologian perusteella. Tätä ei kuitenkaan tule kutsua lääketieteen paradigman muutokseksi, vaan enintään metodologisen paradigman muutokseksi, jos siksikään. Luultavasti parempi ilmaus olisi metodologisen painopisteen siirtyminen kohti RCT:ia. Tämä painopisteen siirtyminen ei kuitenkaan ole synonyymi sanoille ”näyttöön perustuva”.

Voidaankin perustellusti todeta, että lääketiede käsitteenä sisältää jo vaatimukset, joita EBM sille asettaa, minkä vuoksi etuliite ”näyttöön perustuva” on merkityksetön. Sen sijaan tämä metodologinen siirtymä, josta EBM-ajattelussa on kyse, voi hyvin olla lääketieteeseen kuuluvaa ”korkean ammattitaidon ylläpitämistä”. Kuitenkaan tähänastisen tarkastelun perusteella ei voida sulkea pois mahdollisuutta, että tämän siirtymän omaksuminen saattaa rajoittaa lääkäreitä tavalla, jonka vuoksi he eivät voi saavuttaa korkeinta mahdollista ammattitaitoa. Seuraavassa kappaleessa tarkastelen tämän siirtymän oikeutusta, etuja ja haittoja.

4 Näyttöön perustuva lääketiede – hyviä ja huonoja puolia

4.1 Esimerkki EBM:n eduista: rytmihäiriölääkkeet sydämen vajaatoiminnassa

Eräs merkittävistä EBM:n menestystarinoista on systemaattinen katsaus rytmihäiriölääkkeistä eteisvärinän hoidossa. Tuohon saakka oli patofysiologisen päättelyn perusteella ajateltu, että rytmihäiriön hoidossa rytmin korjaaminen on ymmärrettävästi paras tapa hoitaa potilasta. Tutkimus kuitenkin osoitti, että rytmihäiriölääkkeet assosioituivat kohonneeseen sydänkuolleisuuteen ja rytmihäiriökuolleisuuteen. Alaryhmäanalyysissä havaittiin, että tämä löydös selittyi sydämen vajaatoimintaa sairastaneilla potilailla, joilla riski oli merkittävästi kohonnut.¹³

Tätä löydöstä ei ollut mahdollista päätellä sen aikaisen tiedon valossa, vaan assosiaatio löydettiin ensin systemaattisella analyysillä, joka mahdollisti löydöstä selittävien patofysiologisten mallien kehittämisen. Tämä "what works" –ajattelun keskeinen teesi on helposti allekirjoitettavissa: vaikka taustateoria kuinka tukisi jotain selitystä, täytyy silti lopputuloksena olla parempia hoitotuloksia. Kuitenkin on hyvä muistaa, että patofysiologinen päättely ohjaa väistämättä toimintaamme. Tutkimuksia ja systemaattisia katsauksia tehdään vain silloin, kun on odotettavissa jonkinlainen looginen vaste hoidosta. Mikäli jo 'a priori' oletamme, että minkäänlaista eroa ei tulla havaitsemaan, ei tutkimuksen suorittaminen ole tarkoituksenmukaista. Tämä ei tarkoita, että negatiivinen tulos olisi turha, vaan että koeasetelmaa suunniteltaessa tutkijoilla on tyypillisesti jonkinlainen oletamus, jonka todenperäisyyttä he haluavat testata.

Tämä tapausesimerkki on ollut tärkeä todistuskappale pelkän patofysiologisen päättelyn rajallisuudesta ja systemaattisemman analyysin ja tutkimuksen tarpeellisuudesta. Tärkeänä sivuhuomiona tässä tutkimuksessa ei kuitenkaan potilaita ollut satunnaistettu. Käsittelemme satunnaistamisen hyötyjä ja haittoja myöhemmin, mutta jo tässä kohtaa on hyvä huomata, että myös ei-satunnaistettu tutkimus ja alaryhmäanalyysi voivat

tuottaa EBM:n kannattajien mielestä mielekästä tietoa, vaikka hierarkiassa ja käytännön tasolla niiden vastustus on usein voimakasta.

4.2 Tutkimuksen virhelähteiden huomioiminen koasetelmaa laatiessa

Yksi keskeisistä metodologisista kriteereistä, joita EBM:n yhteydessä esitetään, on satunnaistamisen merkitys tutkimuksen laadulle. Jo EBM working groupin artikkelissa 1992 esitetään metodologiset kriteerit tutkimusten laadun arvioimiseksi ja todetaan ”jos tutkimus ei ole satunnaistettu, se on turha, eikä tarjoa arvokasta tietoa.”²

Myöhemmin 1996 Sackett toteaa myös ”koska satunnaistettu koe, ja erityisesti systemaattiset katsaukset useista satunnaistetuista kokeista, ovat paljon todennäköisempiä informoimaan ja huomattavasti epätodennäköisempiä johtamaan meitä harhaan, siitä on tullut ’kultainen standardi’, kun arvioidaan aiheuttaako hoito enemmän hyötyä vai haittaa.”⁴ Kummassakaan artikkelissa ei kuitenkaan käydä läpi tarkemmin, miksi tai mikä satunnaistamisessa johtaa esitettyihin väitteisiin sen tuottaman tiedon laadusta ja satunnaistamattoman tiedon arvottomuudesta.

Satunnaistamisen historiaa sekä hyviä ja huonoja puolia avaa kattavasti Peter Urbach artikkelissaan ”The value of randomization and control in clinical trials”. Hänen mukaansa satunnaistaminen on yksi esitetyistä ratkaisuista kliinisten tutkimusten perusongelmaan: milloin voimme arvioida jonkin hoidon toimineen?¹⁴

Perusajatuksena on, että jos ryhmä A saa hoitoa ja ryhmä B vertailuhoitoa (tai plaseboa) ja ryhmässä A mitattavat hoitotulokset ovat parempia, niin ”muiden asioiden ollessa samanlaiset” (other things being equal), voidaan A:n saaman hoidon sanoa olevan parempaa hoitoa.¹⁴

Itse päättely on pätevää, jolloin keskeiseksi kysymykseksi nousee milloin ”muut asiat ovat samanlaisia”? Usein esitetty, mahdollisesti realistinen tavoite on, että ”ryhmien

tulisi olla samankaltaiset ennustavien tekijöiden (prognostic factors) osalta – siis tekijöiden, jotka vaikuttavat sairauden kulkuun.” Kuitenkin tämän lähestymistavan ongelmana on, että näiden ennustavien tekijöiden kokonaismäärä on meille tuntematon, ja monet niistä eivät koeasetelemaa laatiessa välttämättä nouse mieleen tai niiden virheellisesti ajatellaan olevan merkityksettömiä. Tutkijoille täysin tuntemattoman tekijän eroa ryhmien välillä ei voida ollenkaan mitata, mutta se voi silti aiheuttaa vääristyneitä tuloksia. Urbachin mukaan tämä on kliinisten kokeiden keskeinen ongelma.¹⁴

Urbach esittelee Fisherin ehdotuksen jakaa tekijät tunnettuihin ja tuntemattomiin. Tunnettujen osalta Fisherin mukaan on järkeenkäypää varmistaa ryhmien samankaltaisuus. Esimerkkinä tunnetusta tekijästä on potilaan oma kokemus hoidon vaikuttavuudesta ja sen merkitys hoitotuloksiin, jolloin sokkouttaminen poistaa tämän eron ryhmien väliltä.¹⁵ Tuntemattomien tekijöiden tasaamiseksi Fisherin ajatusten pohjalta on ehdotettu satunnaistamista.¹⁴

Satunnaistamisen puolesta Urbach esittelee Byarin ajatuksen, että ”satunnaistamisella on taipumus tasata vertailuryhmät ... ennustavien tekijöiden osalta, olivatpa nämä tekijät tunnettuja tai tuntemattomia”¹⁶. Kuitenkin Urbach toteaa, että ei ole mitenkään itsestään selvää, etteikö jokin tekijä voisi olla painottunut toisessa ryhmässä, erityisesti kun tuntemattomien tekijöiden kokonaismäärää ei edes voida tietää. Lisäksi Urbach toteaa, että esimerkiksi lääkekapseliin valmistusprosessissa vahingossa päätyneet muu aine voi vaikuttaa hoidon lopputulokseen, mutta tällaisten tekijöiden osalta satunnaistaminen ei tarjoa minkäänlaista suojaa.¹⁴

Artikkelinsa loppupuoliskolla Urbach tarkastelee satunnaistamisen ja muiden tutkimusmetodologisten työkalujen arvoa bayesilaisen ajattelun valossa. Lyhyesti tässä tarkastelussa on kyse tietyn evidenssi E:n olemassaolon oikeuttamisesta jonkin hypoteesin valossa. Kun tutkimus on tuottanut tietyn näytön, tulee arvioida, mikä hypoteeseista on todennäköisin selittämään tämän löydöksen.¹⁴

Tällöin yksittäisen hypoteesin todennäköisyys on kuvattavissa muodossa: kyseisen hypoteesin osatodennäköisyys jaettuna kaikkien mahdollisten hypoteesien yhteenlasketuilla osatodennäköisyyksillä. Urbach antaa tyypistetyn esimerkin, jossa evidenssille on vain kaksi mahdollista selitystä: interventio vaikuttaa positiivisesti tai vaikutus selittyy potilaan uskomuksella hoidon vaikuttavuudesta. Lisäksi oletetaan, että vertailuryhmät ovat merkittävässä määrin samankaltaisia.¹⁴

Jos vain nämä kaksi mallia ovat mahdollisia post hoc –selityksiä evidenssille, on tärkeää laatia koeasetelma siten, että saataessa positiivinen tulos, ”interventio vaikuttaa” -selityksen suhteellinen todennäköisyys on mahdollisimman suuri. Tällöin ”uskomus hoidon vaikutuksesta” –selityksen todennäköisyys kannattaa minimoida. Tähän sokkouttaminen on tehokas menetelmä.¹⁴

Tämä Bayesilainen lähestymistapa tutkimusasetelman laatimiseen on analoginen aiemmin esitellyn evidenssikäsityksen kanssa: ”E vahvistaa H:ta, jos se lisää uskottavasti syytä ottaa H totena eikä samanaikaisesti nosta samassa määrin muita mielekkäitä hypoteeseja.”⁸ Jotta tutkimus antaisi hyödyllistä tietoa, on tutkimusta laadittaessa pyrittävä varmistamaan, että E todella vahvistaa ensisijaisesti H:ta eikä muita hypoteeseja.

Urbach jatkaa myös, että vastaavasti voidaan olettaa kolmas mahdollinen selitys vaikuttavuudelle: lääkäri kykenee kokemuksensa avulla intuitiivisesti havainnoimaan potilaastaan esimerkiksi vastustuskykyä sairaudelle, sairauden vakavuusastetta tai muita hoitoon vaikuttavia tekijöitä ja tiedostamattaan käyttää näitä päättäessään, sijoitetaanko potilas hoito- vai kontrolliryhmään. Tällöin evidenssi ei kuvaisikaan lääkkeen vaikuttavuutta vaan lääkärin kliinisten intuition tasoa. Tämän hypoteesin todennäköisyyden minimoimiseen satunnaistaminen on yksi tehokas keino.¹⁴

Urbach toteaaakin, että ”on selvää, että tutkimuspotilaiden satunnaistaminen on välillä hyödyllinen keino samankaltaistaa tutkimusryhmät, erityisesti jos on syytä epäillä, että lääkärit saattaisivat jakaa potilaat ryhmiin heidän sairautensa vakavuusasteen perusteella. Kuitenkaan satunnaistaminen ei ole absoluuttisen välttämätöntä, se ei ole

välttämätön ehto (latin. sine qua non), se ei ole ainoa, eikä aina edes paras menetelmä ryhmien muodostamiseen kliinisessä kokeessa.”¹⁴

Artikkelinsa lopussa Urbach esittää, että bayesilaisen ajattelun pohjalta paras lähestymistapa klinisiin kokeisiin vaikuttaisi olevan, että asiantuntijat listaavat tekijät, jotka heidän mielestään voivat vaikuttaa taudin kulkuun [sekä potilaan että hoidon osalta] ja tämän jälkeen erillinen henkilö tai paneeli muotoilee asianmukaisen koeasetelman, jossa listatut sudenkuopat voidaan välttää. Urbach esittää myös toiveen, että satunnaistamisen välttämättömyydestä voitaisiin luopua ja käyttää sitä vain asianmukaisiin käyttöaiheisiin. Tällöin esimerkiksi retrospektiivisten kontrolliryhmien käyttö voisi muuttua mahdolliseksi, ja jos tämän koeasetelman ongelmat kyettäisiin ratkomaan, avaisi se oven eettisemmälle tutkimukselle, jossa osaa potilaista ei tarvitsisi altistaa huonolaatuisemmaksi uskotulle hoidolle. Urbachin mielestä suuri ongelma on, että satunnaistamisen vaatimus on niin suuri, että edes keskustelua aiheesta ei voida käydä, vaikka sen tuloksilla voisi olla huomattavaa merkitystä lääketieteen tutkimukselle.¹⁴

4.3 Satunnaistettu kontrollitutkimus – laadukkainta tietoa?

4.3.1 Mistä tutkimushierarkia nousee ja mikä näyttö tukee tätä hierarkiaa?

Evidenssihierarkioita on esitetty EBM-keskustelun aikana useita toisistaan hieman eroavia vaihtoehtoja, mutta yhtenäisenä linjana on, että RCT:it ovat hierarkian yläpäässä ja anekdotaaliset havainnot vähiten arvostettua dataa. Kuten Bluhm toteaa, kyseessä ei kuitenkaan tarkalleen ottaen ole evidenssin hierarkia vaan tutkimusmetodologioiden välinen hierarkia¹⁷.

Oxman et al. esittelivät GRADE-järjestelmän näytön laadun arvioimiseksi hoitosuosituksia varten. Järjestelmässä tutkimuksen painoarvo määräytyy aluksi sen metodologian perusteella, jonka jälkeen painoarvoon voi tulla muutoksia, jos esimerkiksi tutkimuksen toteutus on ollut heikkolaatuinen tai löydös on tilastollisesti poikkeuksellisen merkittävä. Tutkimusasetelmat on karkeasti jaettu havainnoiviin ja satunnaistettuihin tutkimuksiin. Kirjoittajat esittävät, että niin loogiset argumentit kuin empiirinen näyttö tukevat tätä luokittelua. Esimerkkinä he esittävät havainnoivien tutkimusten löydöksen, että hormonikorvaushoito vähentäisi sepelvaltimotaudin riskiä, mutta satunnaistetut kokeet löysivät assosiaatiota jopa kohonneeseen riskiin.¹⁸

Kuitenkin tämä esimerkki metodologioiden eroista on yksittäinen, anekdotaalinen havainto, eikä ”parasta mahdollista tutkimusnäyttöä”. Kuten Norman toteaa: ”vieläkään emme ole saaneet näyttöä, joka tukisi väitettä EBM:n tehokkuudesta, paitsi yhden kohorttitutkimuksen, jossa vertaillaan McMasterin ja Toronton opiskelijoita. Vaikka tätä tutkimusta mainostetaan parhaana tarjolla olevana näyttönä EBM:n käytännön eroista, todellisuudessa havaitut erot perustuivat monivalintakokeeseen, eivät kliiniseen suoriutumiseen, minkä lisäksi erot tasoittuivat valmistumisen jälkeen. ... Meillä ei ole suoraa, EBM-kelpoista näyttöä, että EBM on reitti totuuteen. Sen sijaan, EBM:n puolustajat tuottavat tapausselostuksia RCT:eiden ja metodologisesti alempiarvoisten tutkimusten eroista.”³

Siis keskeisimmän päätetapahtuman, paremman hoidon, osalta ei ole olemassa EBM-kelpoista näyttöä, että EBM-kelpoisen tutkimuksen tekeminen ja seuraaminen hoitopäätöksissä johtaisi parempaan lopputulemaan.

Toinen keskeinen kritiikki koskee eri tutkimusmenetelmien tulosten välisiä ristiriitoja. EBM:n puolestapuhujat ovat toistuvasti esittäneet, että ei-satunnaistettu koeasetelma on turha, minkä tueksi on esitetty, kuinka havainnoivat tutkimukset ja satunnaistetut tutkimukset saavat keskenään erilaisia tuloksia. Grossman¹⁹ kuitenkin toteaa, että tutkimuksen laadukkuuteen vaikuttaa huomattava määrä muitakin tekijöitä eikä pelkkä satunnaistaminen itseisarvoisesti aseta tiettyä tutkimusta parempilaatuiseksi ja vertaa tällaisen oletuksen tekemistä argumenttiin, jonka mukaan ”koska pituus on tärkeää hyvälle korkeushyppäjälle, täytyy 1,83 m pitkän humalaisen ja selkäydinvaurioista kärsivän olla parempi hyppääjä kuin 1,8 m pitkän olympiaurheilijan.”

Grossman toteaa myös, että eriävien tulosten selittäminen johtopäätöksellä, että RCT:in antama tulos on oikea, on pohjimmiltaan kehäpäätelmä: väitettä ei voi tehdä olettamatta jo etukäteen juuri RCT:in antavan oikean tuloksen.¹⁹

Borgerson kyseenalaistaa koko ajatuksen tutkimusmenetelmien hierarkiasta todeten, ettei sellainen ole millään lailla intuitiivinen tai tieteentekijöiden yleisesti harjoittama käytäntö. Hänen keskeinen kritiikkinsä koskeekin sitä, että mitään hierarkiaa ei ole puolustettu eksplisiittisesti.⁷ Hänen kritiikkiään RCT:ien taipumuksesta suosia probabilistisia kausaation syitä yli mekanististen käsittelen tarkemmin luvussa 3.1.

Kysymykseen, miksi RCT:in antama data olisi yksikäsitteisesti parempaa ja luotettavampaa kuin muiden tutkimusasetelmien, ei ole saatu hyvää vastausta, mutta tätä oletusta kohtaan on esitetty runsaasti laadukasta kritiikkiä.

4.3.2 RCT ja tutkimusmetodologioiden ennakkotodennäköisyydet

Arvioitaessa erilaisia tutkimusasetelmia, on tärkeää ymmärtää niiden erilainen rooli riippuen tutkimuksen tavoitteesta. Vandenbroucke esittelee esseessään "Observational research, randomised trials, and two views of medical science" kaksi eri lähestymistapaa lääketieteen tutkimukseen ja vertailee niiden eroja.²⁰

Ensimmäinen lähestymistapa on havainnoiva, jossa uutta tietoa saadaan, kun asiat nähdään aiemmasta poikkeavalla tavalla: sairaus etenee odottamattomasti tai saadaan selittämätön havainto tai poikkeama alaryhmässä. Tällöin tärkeässä roolissa on uuden oivalluksen todenperäisyyden kartoittaminen: hypoteesi pyritään varmentamaan karkeasti ja julkaisun jälkeen muut tutkijat voivat kokeilla hypoteesin pitävyyttä omassa (jo kerätyssä) datassaan tai kehittää uusia tutkimusasetelmia tutkimuksen eteenpäin viemiseksi.²⁰

Arvioivassa lähestymistavassa taas pyritään selvittämään, kuinka tehokas tietty hoito oikeasti on. Tällöin keskeisessä roolissa on tutkimusasetelman huolellinen suunnittelu ja etukäteen julkistetut kriteerit. Tällöin Vandenbroucken mukaan nähdään, ettei ennakkoon asetetuista kriteereistä tulisi poiketa, vaikka datasta löytyisi lupaavan näköisiä "sivulöydöksiä".²⁰

Vandenbroucke esittää, että arvioivassa lähestymistavassa "evidenssihierarkian" huipulla on RCT ja pohjalla anekdotaalinen evidenssi (yksittäiset potilastapaukset jne.) kun taas observoivassa lähestymisessä hierarkia on päinvastainen. Tämä johtuu erilaisista tavoitteista: observoivassa tavassa tärkeitä ovat uusien ajatusten ja yksittäistapausten nostamat ajatukset uudesta selitysmallista tai hypoteesista. Arvioivassa lähestymisessä taas pyritään selvittämään tietyn yksittäisen tekijän vaikuttavuutta, jolloin muiden vaikutusmekanismien poissulku on välttämätöntä.²⁰

Tämän poissulkemisen merkitys näkyy hyvin hoitoindikaatioiden aiheuttamassa harhassa. Jos oletamme tietyn potilaan saavan hyötyä hoidosta, määräämme sitä hänelle. Vastaavasti mitä sairaampi potilas on, sitä aggressiivisempi lääketieteellinen

interventio tyypillisesti on. Tällöin indikaatio ja potilaan vointi itsessään säätelevät ennustettavissa olevaa hoitovastetta. Hoidon vaikuttavuuden arvioimiseksi satunnaistaminen on usein tehokkain tapa välttää tällainen indikaatioharha (confounding by indication).²⁰

Esseessä esitetään myös, että tieteellisissä piireissä usein intuitiivisesti ajatellaan ennen datan näkemistä päätettyjen alaryhmäanalyysien olevan luotettavampia kuin datan näkemisen jälkeen linjattujen. Kuitenkin Vandenbroucke toteaa, että ”yleisellä tasolla tällaiselle väitteelle on vaikea löytää loogisia todisteita, koska usein tieteen uudet ideat saavat uskottavuutensa siitä, että ne kykenevät selittämään aiempia löydöksiä, joita ei ennen ole kyetty ymmärtämään.”²⁰ Kuten Kvernbekk Newtonin gravitaatioteoriasta puhuessaan totesi, keskeinen kysymys lienee ollut: ”mikä näkymätön tekijä selittää tällaiset löydökset?”⁵ Gravitaatioteoriaa ei voida pitää asianmukaisesti laadittuna, jos vasta havaintojen jälkeen laadittuja selitysmalleja pidetään ehdottoman kiellettyinä. Lisäksi gravitaatiomalli kykeni nimenomaan Vandenbroucken ajatuksen mukaisesti selittämään taivaankappaleiden keskinäisiä liikkeitä ja niissä jo aiemmin ihmeteltyjä poikkeavuuksia.

Usein RCT:in tapauksessa kuitenkin tarkoin suunniteltu ja harkittu koeasetelma tukee enemmän tätä etukäteen ilmoitettujen kriteerien seuraamista. Koska kokeet perustuvat ”vuosien perusteelliseen asiantuntijaharkintaan”, on epätodennäköistä, että jokin merkittävä alaryhmä jäisi huomioimatta etukäteen. Todennäköisempää on, että tällaiset sekoittavat tekijät on jo huolellisesti rajattu tutkimuksen ulkopuolelle. Tällainen ennakkorajaus nostaa todennäköisyyttä, että kokeen tulos on luotettava, mutta vastaavasti laskee todennäköisyyttä sille, että yllättävä poikkeava löydös johtuisi mistään muusta kuin sattumasta. Tällaisen sattuman minimoimiseen taas suuri n-luku on varmin metodi.²⁰

Observoivassa lähestymistavassa puolestaan erikoislaatuisen löydöksen keskeisin aiheuttaja ei ole sattuma, vaan koeasetelman luomat vääristymät ja harhat. Tällöin tutkimuksen toistaminen identtisenä säilyttää samat harhat eikä hypoteesin luotettavuudelle saada lisätukea. Tämän vuoksi Vandenbroucken mukaan tärkeää

olisikin tehdä lukuisia kokeita, joissa astelemaa ja menetelmiä varioitaisiin.²⁰ Jos näissä tarkasteluissa nähdään toistuvia trendejä, nousee luonnollisesti hypoteesin ennakkotodennäköisyys.

Tällöin, jos metodologia ja hypoteesi ovat yhteensopivia, voidaan ilmiötä asiantuntijoiden harkinnan jälkeen tutkia RCT:illa. Oikean positiivisen löydöksen ennakkotodennäköisyys on kuitenkin tässä kohtaa jo huomattavasti suurempi kuin alun perin yksittäistä anekdotaalista potilastapausta kuvatessa. Vandenbroucke kiteyttääkin, että ”olemme saattaneet johtaa itseämme harhaan [RCT:in] uniikista ylivertaisuudesta, koska niillä on alun perinkin huomattavasti korkeampi ennakkotodennäköisyys kuin observeivalla tutkimuksella”²⁰.

Tämän korkeamman ennakkotodennäköisyyden vuoksi RCT:ilta voidaan hyväksyä myös vähemmän merkittäviä löydöksiä ($p < 0.05$). Jos RCT:illa pyrittäisiin tutkimaan teoriaa, jonka ennakkotodennäköisyys on $1 / 100\,000$, olisi positiivinen löydös todennäköisimmin sattumaa kuin oikea positiivinen²⁰, koska $1 / 20$ ($p < 0.05$) on todennäköisyyksissä selvästi isommassa kertaluokassa kuin $1 / 100\,000$ ($p < 0.00001$).

Tätä Vandenbroucken ajatusta voidaan kehittää edelleen soveltamalla bayesilaista lähestymistapaa. Ajatellaan tilannetta, jossa ennakkotodennäköisyys teorian paikkansapitävyydelle on $1/20$ ja tilastollisesti merkittävän löydöksen rajana on $p < 0.05$. Tällöin saadessamme positiivisen löydöksen, sen selityksinä sattumalöydös ($1/20$) ja hypoteesin paikkansapitävyys ($1/20$) olisivat yhtä todennäköisiä. Jos tällä p-arvolla haluamme, että sattuman rooli löydöksen selittäjänä jää suhteessa riittävän pieneksi, täytyy hypoteesin mukaan positiivisen löydöksen ’a priori’ -todennäköisyyden olla huomattavasti enemmän kuin $1/20$.

Vandenbroucken keskeinen teesi onkin, että tarvitsemme monenlaisia tutkimusasetelmia tieteen tekemiseen. Havainnoivalla ja uutta tietoa tuottavalla tutkimuksella on tärkeä rooli, jossa anekdotaalisetkin havainnot nostavat hypoteesin todennäköisyyttä. Kun jokin hypoteesi on riittävän todennäköinen tai nousee sellaiseksi

havainnoivan tutkimuksen avulla, on RCT tärkeä ja hyödyllinen työkalu todellisen vaikutuksen olemassaolon arvioimiseksi.²⁰

4.3.3 RCT:in suosimisen ongelmat

Olen nyt käsitellyt hierarkian ongelmia sekä eri tutkimusmenetelmien käyttötarkoituksia. Tarkastelun perusteella voidaan sanoa, että RCT:in asema on ylikorostunut lääketieteellisessä keskustelussa ja päätöksenteossa: ”julkaisut haluavat RCT:eita, koska he ja heidän lukijansa ovat päätyneet pitämään niitä korkeimman laatuksena.”²¹ Kuitenkin käytännön tasolla tämä ilmiö johtaa ongelmiin vain, jos on olemassa vaara, että RCT:in suosiminen heikentää hoidon laatua.

Valitettavasti RCT:in nostaminen yli muiden tutkimusasetelmien on vaarallista kahdesta syystä: toisaalta sen luotettavuus ei ole läheskään niin itsestään selvä kuin millaisena sitä käytännössä pidetään ja toisaalta on vaarana, että hoitomuodot yksipuolistuvat.

Ensimmäinen ongelma liittyy vahvasti tutkimusrahoitukseen. Jos tutkimuksella on kaupallista potentiaalia, sillä on huomattavasti parempi mahdollisuus saada yrityksiltä rahoitusta²². Tämä johtaa tutkimusaiheiden valikoitumiseen kaupallisen potentiaalın perusteella todellisen tarpeen sijaan, mitä voidaan syystäkin pitää huolestuttavana. Vain voittoa tarkastellessa parantavat interventiot ovat selvästi huonompia vaihtoehtoja kuin pitkäaikaissairauksien oireenmukaiset tai ennustetta parantavat interventiot, joita potilas joutuu todennäköisesti käyttämään lopun ikäänsä.

Tämän lisäksi kaupallisten toimijoiden rahoittamat tutkimukset ovat muita tutkimuksia todennäköisemmin positiivisia²². Tämä voi toki johtua myös kaupallisten toimijoiden taidosta valita rahoittamansa tutkimukset keskimääräistä tarkoituksenmukaisemmin, mutta ellemme voi poissulkea mahdollisuutta, että taloudellinen intressi saattaa ruokkia halua saada tutkimuksesta positiivinen löydös, vähentää tämä huomattavasti kaupallisesti rahoitettujen RCT:iden luotettavuutta. Tämän lisäksi RCT:issa voi olla systemaattisia tai rakenteellisia virheitä, joita ei tutkimusasetelmaa laatiessa ole kyetty

ottamaan huomioon. Näitä olen tarkastellut tarkemmin luvussa ”Tutkimuksen virhelähteet koeasetelmaa laatiessa”.

Gupta kuvaa artikkelissaan ”A critical appraisal of evidence-based medicine: some ethical considerations” myös teknisen harhan ja julkaisuharhan aiheuttamia ongelmia.²³ Tekninen harha ilmenee siten, että ymmärrettävästi meillä on taipumus laatia sellaista tutkimusta, jota osaamme tehdä, minkä vuoksi tutkimme enemmän aiheita, joita osaamme tutkia. Tämä luo systemaattisen harhan, joka vaikuttaa siihen, millaista tietoa ylipäättään tuotamme. Jos suosimme RCT:ia koeasetelmana yli muiden, saamme laadukasta näyttöä vain hoidoista, joita RCT:illa voi tutkia ja muut hoidot jäävät tällöin paitsioon.²³ ”Tämä tarkoittaa, että EBM on pitkälti kääntynyt pois kysymyksistä ja kokemuksista, joita on vaikea määritellä, mitata tai kategorioida, kuten elämänlaatuun ja kuolemaan liittyvät kokemukset, sekä äitiyteen ja seksuaalisuuteen liittyvät kokemukset.”²¹ Pahimmassa tapauksessa saamme myös huonolaatuista näyttöä hoidoista, joita on yritetty tutkia RCT:illa, vaikka se ei ollenkaan soveltuisi kyseisen aiheen tutkimiseen.

Näiden lisäksi julkaisuharha vääristää entisestään RCT:in luotettavuutta, kun negatiiviset tutkimukset jätetään julkaisematta. Gupta arvioi, että positiivisia tuloksia suositaan, koska tämä tekee lääketieteellisestä kirjallisuudesta kiinnostavampaa. Kuitenkin tämä voi luoda harhaanjohtavan kuvan uuden intervention tehosta, kun keskeinen negatiivinen data ei ole yhtä lailla saatavilla. Lisäksi Gupta toteaa, että julkaisuharha voi johtaa tekniseen harhaan ja tiettyjen tutkimusaiheiden jäämiseen liian vähälle huomiolle, kun tutkijat valitsevat aiheita, joista saadaan helpommin julkaisijoille kelpaavia tuloksia. Esimerkiksi psykiatriassa RCT:it korkealle asettava hierarkia johtaa tekniseen harhaan, jossa farmakologisia hoitomuotoja tutkitaan todennäköisemmin kuin psykoterapeuttisia hoitomuotoja.²³

RCT:illa on siis sekä mahdollisuus olla väärässä että nostaa tietyt hoitomuodot suurempaan suosioon, kun niiden tueksi on saatavissa ”laadukkaampaa” näyttöä. Lääkärin velvollisuus on kuitenkin antaa potilaille parasta mahdollista hoitoa, ei

laadukkaimpaan näyttöön perustuvaa. Satunnaistetulla tutkimuksella on ehdottomasti paikkansa erityisesti lääkkeellisten interventioiden vaikuttavuuden tutkimisessa, mutta sen nostaminen yksiselitteisesti yli muiden on sekä metodologisesti kestänyt että johtaa helposti epätarkoituksenmukaiseen hoitoon.

4.4 Retroaktiivisen rukouksen vaikutus – taustateorian ja tutkimuksen välisestä yhteydestä

Vuonna 2001 Leibovici julkisti BMJ:n joulunumerossa humoristisen tutkimuksen retroaktiivisen rukouksen vaikutuksista sepsis-potilailla. Tutkimukseen otettiin mukaan kaikki 1990-96 välillä sairaalassa olleet bakteeriviljelyllä todennettua sepsistä sairastaneet potilaat. Potilaat jaettiin satunnaisesti kahteen ryhmään, joiden välillä ei huomattu eroja perustiedoissa (baseline characteristics). Tämän jälkeen hoitoryhmä arvottiin heittämällä kolikkoa ja hoitoryhmän etunimet annettiin henkilölle, joka rukoili lyhyesti ryhmän puolesta. Sekä potilaat että hoitohenkilökunta olivat täysin sokkoutettuja intervention suhteen. Kuolleisuudessa ei nähty merkittävää eroa, mutta sairaala-ajan ($p < 0.01$) ja kuumeen keston ($p < 0.04$) suhteen hoitoryhmässä havaittiin merkittävä etu.²⁴

Kuten Worrall (2010) toteaa, ”on sanomattakin selvää, että tätä [tutkimusta] ei tule ottaa vakavasti”²⁵. Jo intuitiivisesti huomaamme, että emme pidä tätä tutkimusta uskottavana, vaikka se metodologisesti täyttää kaikki laadukkaan RCT:n kriteerit. Johdonmukaisesti (vaikkakin satiirisesti) Leibovici toteaa tutkimuksessaan ”Toistaiseksi ei tunneta [patofysiologista tai biologista] mekanismia, jolla etäältä lausuttu retroaktiivinen rukous vaikuttaisi. Kuitenkin [tilastollisesti] merkittävä tulos ja [tutkimuksen] virheetön toteutus todistavat, että vaikutus saatiin aikaan.”²⁴ Hän jatkaa kritiikkiään ja noudattaen EBM:n teesejä - 1. Päätösten tulee perustua parhaaseen mahdolliseen tietoon sekä 2. RCT tarjoaa parasta mahdollista tietoa - päättyy ainoaan mahdolliseen johtopäätökseen ”Interventio on kustannustehokas ja sitä tulisi harkita kliiniseen käyttöön”²⁴. Leibovici on itse myös kirjoittanut ”jos vaikutusmekanismin

todennäköisyys[/uskottavuus] on äärettömän pieni, ei koetulos kykene kääntämään tätä todennäköisyyttä eikä koetta tulisi tehdä”.²⁶

Giacomini käsittelee artikkelissaan sekä Leibovicin katsausta että yleisemmin rukouksen tutkimista RCT:lla ja päätyy lopputulokseen, että metodologisesti menetelmä ei ole mielekäs rukouksen tutkimiseen. Hän esittää, että ”lääketiede – myös näyttöön perustuva lääketiede – on teorioihin perustuvaa. Teoria antaa datalle sen statuksen fakteina, ohjaa uusien faktojen etsimistä ja tarjoaa löydöksille niiden tarkoituksen ja implikaatiot.”²⁷

Retroaktiivisen rukouksen tutkimuksesta Giacomini sanookin, että ”ei ole esimerkillistä tiedettä – mutta edustaa huonoa etiikkaa – tutkia black box interventiota, jonka kausaalisuhdetta terveyteen ja haittoihin ei voida käsittää. Jotta sitä voitaisiin arvioida tieteellisesti, tulee hoidon taustalla olevan perustelun olla läpinäkyvä.”²⁷

Giacomini käy läpi rukoustutkimuksen keskeisiä yhteensopivuusongelmia RCT:in ja taustaoletuksena olevan persoonallisen, omnipotentin Jumalan välillä. Koska jokaisen sielun yksilöllinen suhde Jumalaan on merkityksellinen, ovat testipotilaat (ja rukoilijat) uniikkeja eivätkä edusta populaatiota. Näin ollen tulosten yleistäminen on taustateorian vuoksi täysin mahdotonta, vaikka valintakriteerit olisivat laadukkaat. Lisäksi persoonallisen Jumalan vastaukset rukouksiin ovat aina uniikkeja, joten niiden mallintaminen luonnonlakina tai kausaalisuhteena on mahdotonta. Lisäksi on täysin mahdotonta suojella kontrolliryhmää omnipotentin Jumalan interventiolta, jos hän sellaisen tahtoo tehdä.²⁷

Giacomini mukaan rukouksen parantavan vaikutuksen tutkimusten tarkastelu osoittaa rakenteellisia sokeita pisteitä nykyisen EBM:n kyyvyssä arvioida näyttöä ja tarjoaa varoittavan esimerkin siitä, mitä voi tapahtua, jos tutkimme jotain, minkä vaikutusmekanismia emme ymmärrä, mikä Giacominin mukaan on suosittu idea EBM-orientoituneiden tutkijoiden keskuudessa. Hän vetää johtopäätöksen ”hyvän kausaalisen argumentin puuttuessa, tulisi tuloksiin suhtautua skeptisesti.”²⁷

Tämän tarkastelun perusteella voidaan tehdä johtopäätös, että tutkimuksen ja metodologian tulee olla tarkoituksenmukainen huomioiden taustateoria vaikutusmekanismista ja että RCT:in tarjoama positiivinen tulos on merkityksetön, jos sen selitykseksi ei ole tarjottavissa uskottavaa teoreettista mallia vaikutuksesta.

4.5 Juridinen tapauselostus: PSA-seulonta ja hoitolinjat

Länsimaisessa yhteiskunnassa hyväksyttävät tavat toimia määrittää viime kädessä oikeuslaitos. Tästä traagisen esimerkin esittelee Daniel Merenstein. Hän kuvaa 53-vuotiaan potilaan, jolle hän on lääkärintarkastuksen yhteydessä kertonut PSA-testauksen hyödyistä, riskeistä ja niiden välisestä tasapainosta. Tämän tiedon perusteella potilas päätti yhteisymmärryksessä lääkärin kanssa pidättäytyä PSA-tutkimuksesta.²⁸

Myöhemmin sama potilas kävi eri lääkäriä, joka tilasi PSA:n keskustelematta asiasta potilaan kanssa. Arvo oli korkea ja potilaalta löydettiin pitkälle kehittynyt prostatasyöpä ilman hoitomahdollisuuksia (Gleason 8). Merenstein kuvaa, että tällainen syöpä on hyvin vaikea löytää varhaisessa vaiheessa eikä lääketieteellinen kirjallisuus tue ajatusta, että aikaisempi diagnoosi olisi muuttanut potilaan ennustetta.²⁸ Tämä intuitiivinen halu uskoa että ”jotain olisi ollut tehtävissä” vaikka tilastot sanovat selvästi päinvastaista on mielestäni hyvä esimerkki vividness-hypoteesista, jota tarkastelen seuraavassa kappaleessa.

Merenstein oli tehnyt tarkat kirjaukset potilaan kanssa keskustelusta ja hänen valitsemaansa linjaa tukivat kaikki kansallisesti tunnustetut lääketieteelliset ryhmittymät sekä kirjallisuus. Hänen yllätyksekseen potilaan lakimiehen argumentaatio perustui siihen, että lääkärin ei olisi tullut keskustella potilaan kanssa vaan tilata tutkimus kysymättä. Syyttäjä myös vetosi siihen, että käytäntö jättää PSA tutkimatta oli vastoin osavaltion yleisiä hoitostandardeja. Syyttäjällä oli neljä lääkäriä, jotka todistivat ottavansa PSA:n rutiininomaisesti kaikilta yli 50-vuotialta miehiltä. Tätä syyttäjän argumentaatiota tarkastellessa on tärkeää muistaa, että uuden tiedon siirtymisessä

hoitokäytänteisiin on viiveitä, joten se, mitä tällä hetkellä tehdään ei välttämättä ole parasta mahdollista hoitoa.²⁸

Syyttäjä vetosi päätöslauselmassaan valamiehistöön, jotta he antaisivat ennakkotapauksen EBM:n harjoittamista vastaan. Valamiehet vapauttivat Merensteinin, mutta hänen sairaalansa tuomittiin miljoonan dollarin korvauksiin. Tämä oli Merensteinin mukaan viesti siitä, ettei valamiehistö luota kansallisiin hoitosuosituksiin tai potilaan mukaan ottamiseen päätöksenteossa.²⁸

Merenstein myös toteaa, että ”Hoitovirhe[selvittelyä] pidetään mekanismina, joka varmistaa, että lääkäreillä on velvollisuus ylläpitää hoidon laatua. Kuitenkin tämä tapaus osoittaa sen aiheuttavan päinvastaista: rankaisevan näytön laittamisesta käytäntöön, mikä hidastaa hoitojen kehittymistä ja kannustaa potilaille huonoihin käytäntöihin.”²⁸

Merensteinin artikkelista on tehtävissä johtopäätös, että vaikka pitäisimme EBM:n esittämiä tapoja parhaina medisiinan harjoittamiseksi, ei ole itsestään selvää, että oikeuslaitos tukee tätä lähestymistapaa. Luotettava tilasto ei välttämättä riitä todisteeksi ihmisten mielissä, jos vastassa on voimakas tarina.

5 Psykologinen näkökulma: Vividness-hypoteesi

Olen nyt esittänyt, että evidenssin käsite ei tue yksittäisen tutkimusmetodologian nostamista ylitse muiden. Myöskään satunnaistaminen ei ole automaattisesti parempi kuin ei-satunnaistettu koeasetelma, vaan päinvastoin se voi joskus olla suorastaan vaarallinen, jos sen oikeuteen luotetaan sokeasti. Lisäksi olen esittänyt, että kahden tapahtuman välisen assosiaation tutkiminen ilman uskottavaa teoriaa niiden välisestä vaikutusmekanismista ei ole tieteellistä. Toisaalta edellisessä kappaleessa esitelty tapaus osoittaa, että laadukkaan tiedon käyttämistä hoidon perusteena ei välttämättä nähdä yhteiskunnallisella tasolla hyväksyttävänä.

Tätä tarkastelua vasten nousee ymmärrettävästi kaksi kysymystä: mikä selittää EBM:n ongelmakohtiin suhteutettuna hämmentävän suuren suosion ja mikä aiheuttaa sen, että luotettavaan tutkimukseen nojaaminen ei juridisesta näkökulmasta näyttäydä asianmukaisena lääketieteenä?

Kiinnostavan näkökulman kysymykseen tarjoaa Normanin artikkeli²⁹, jossa hän referoi psykologian toistuvaa tutkimustulosta, jonka mukaan ihmiset ovat poikkeuksellisen huonoja prosessoimaan tilastollista tietoa ja suorastaan rutiininomaisesti hylkäävät tällaisen informaation, jos heille tarjotaan eläväinen (vivid) vastaesimerkki³⁰. Tilastollinen tieto on juuri se muoto, jossa suurin osa EBM:n suosimasta evidenssistä saadaan.²⁹ Tämä niin kutsuttu vividness -hypoteesi esittää, että eloisa informaatio säilötään ja muistetaan helpommin kuin ”lakoninen” tieto.

Norman esittää, että tämä hypoteesi selittää hyvin EBM-liikehännän vakuuttavuutta, koska EBM itsessään on eläväinen tarina. Meistä tuntuu vaistonvaraisesti järkeenkäyvältä, että lääkärit, jotka pysyvät ajan tasalla alan kirjallisuudesta, kykenevät antamaan parempaa hoitoa. Norman kysyykin retorisesti, ”voiko näin itsestään selvästi tosi oletus tarvita numeerista dataa tuekseen? ... EBM, jos ei mitään muuta olekaan tehnyt, on varmasti tehnyt meistä kollektiivisesti varovaisia hyväksymään todeksi osoittamattomia väitteitä, huolimatta siitä kuinka uskottavilta ne meistä kuulostavatkaan. Valitettavasti EBM:stä esitetyt väitteet kuuluvat tähän samaan

luokkaan.” Vividness -hypoteesi selittää Normanin mukaan myös sen, miksi monet lääkärit ovat vastahakoisia hylkäämään itse tekemiänsä kliinisiä havaintoja niitä vastaan esitetyn luotettavan tilaston kohdatessaan.²⁹

Vividness -hypoteesi selittää erinomaisesti ihmiselle luonteenomaista toimintaa. Esimerkiksi rokotusten vastustajien taipumus nostaa yksittäinen tarina lukemattomien pelastettujen henkien yli näyttäytyy ymmärrettävässä valossa. EBM-keskustelussa taas olisi tarpeellista ymmärtää tilastollisen tutkimustiedon tärkeä tehtävä nykyisten käsitystemme muokkaamisessa ja samanaikaisesti ymmärtää EBM:n olevan rajallinen malli, jonka luotettavuus nykytiedon valossa perustuu intuitioon ja tarinaan eikä luotettavaan näyttöön. Vaikka vividness -hypoteesia on tutkittu vain tilastollisen tiedon osalta, on helppo kuvitella ilmiön toistuvan, kun vastakkain asetetaan elävä tarina ja sitä kritisoiva laadukas, mutta raskas filosofinen argumentaatioketju.

Tärkeä kysymys onkin, missä muissa asioissa sorsumme huomaamattamme seuraamaan yksittäisen tarinan voimaa, vaikkei väitteellä olisi todellisuuspohjaa? Ja kuinka osaisimme tarkastella itsestään selvinä pitämiämme asioita kriittisesti? Omia sokeita pisteitämme emme osaa tunnistaa ja juuri ne asiat, joita emme huomaa, ovat todennäköisimmin niitä, jotka eniten vaatisivat huomiotamme.

6 Mitä EBM:n tilalle?

6.1 Hierarkiaa laajempaa tutkimusten arviointia

Jos ehdoton hierarkia ei ole toimiva ratkaisu kliinisessä lääketieteessä, miten tutkimuksen tarjoamaa dataa sitten tulisi arvottaa ja kuinka tehdä käytännön potilastyötä?

Datan arviointiin erinomaisia ajatuksia on koonnut Robyn Bluhm. Hän esittää, että kokonaisvaltaisempi tutkimusasetelmien tarkastelu voisi auttaa tulkitsemaan ja soveltamaan dataa paremmin. ”Keskimääräinen vaikutus” hoitoryhmässä voi johtua esimerkiksi tulosten suuresta aggregaatiosta keskiarvon ympärillä tai vaihtoehtoisesti pelkästään tietyn alaryhmän saamasta poikkeuksellisen voimakkaasta vasteesta.¹⁷ Tämä ei ole ristiriidassa aiemmin esitellyn Vandenbroucken ajattelun²⁰ kanssa, sillä jos jokin keskeinen ennusteeseen vaikuttava tekijä on jäänyt huomioimatta koeasetelmaa suunnitellessa, se luonnollisesti aiheuttaa toistuvan poikkeaman RCT-tutkimuksissa, jonka havaitseminen vaatii nimenomaan post hoc -alaryhmäanalyysiä.

Bluhm esittääkin, että RCT:in metodologista ehdottomuutta tulisi tietyissä tilanteissa höllentää erityisesti post hoc -tulkintojen osalta. Näitä täydentää luontevasti patofysiologinen ajattelu, jonka tuoma mekanistinen ymmärrys voi tarjota selityksen poikkeavalle datalle.¹⁷ Tällöin toki nousee uusia heikkouksia, kuten riski ajaa alaryhmäanalyyseja, kunnes jotain positiivista löydetään, mikäli tutkimusten valvonta ei ole riittävää. Kuitenkin näkisin, että suurempi vaara on olla tunnustamatta nykyisen mallin heikkouksia, jolloin niihin ei voida etsiä ratkaisuja. Hyödyntämällä yhtäaikaaisesti probabilistisia ja mekanistisia tutkimusmenetelmiä voidaan hyödyntää kummankin vahvuuksia ja välttää yksisilmäisyyden aiheuttamia sudenkuoppia.

Bluhm myös näkee, että kliinisissä tutkimuksissa hierarkiaa mielekkäämpi vaihtoehto olisi tutkimusten jaottelu liukuvalla jatkumolle ”puhtaista” tutkimuksista ”pragmaattisiin” tutkimuksiin. Puhtaiden tutkimusten ominaispiirteenä on suuri sisäinen validiteetti, joka nousee tiukoista potilaiden valintakriteereistä ja intervention

positiivisen vaikutuksen arvioinnista. Samanaikaisesti näiden valintojen seurauksena tunnustetaan, että tulokset ovat vähemmän yleistettävissä yksittäiseen potilaaseen.³¹

Pragmaattiset tutkimukset taas arvioivat kokonaisvaltaisemmin potilaiden saamaa vastetta. Valikoimiskriteerit voivat sekä perussairauksien että muiden hoitojen osalta olla laajemmat, jolloin negatiivisia vaikutuksia ja erityisesti interventioiden vaikutuksia todellisilla potilailla kyetään arvioimaan paremmin. Tällöin kuitenkin tietyn yksittäisen intervention vaikuttavuuden arviointi on vähemmän luotettavaa.³¹

Keskeisenä väitteenä Bluhmilla on, että evidenssihierarkiat eivät puhu tästä puhtas-pragmaattinen –jaottelusta ollenkaan, mutta nykyinen hierarkia on tästä huolimatta taipuvainen suosimaan puhtaita tutkimuksia. Tämän vuoksi Bluhm ehdottaakin siirtymistä kokonaisvaltaisempaan tutkimusten arviointiin, jossa puhtas-pragmaattinen –akselin kaltaiset tyypitykset tarjoavat evidenssihierarkiaa paremman työkalun ymmärtää yksittäisen tutkimuksen vahvuuksia, heikkouksia ja sovellettavuutta kliiniseen työhön.³¹

6.2 Kasuistinen lääketiede

Kaikesta metodologisesta debatista huolimatta lääkärin keskeisin jokapäiväinen kysymys on ”kuinka hoidan tätä potilasta?” Riippumatta tutkimusdatan määrästä, ”on –pitäisi” –välin kurominen jää aina yksittäisen lääkärin vastuulle eikä pelkkien Cochrane-katsauksen automaattinen implementointi jokaiselle potilaalle näytä tarjoavan tyydyttävää vastausta.

Tonelli ehdottaa ratkaisuksi kasuistista eli tapauslähtöistä lähestymistapaa. Tällöin lähtökohtana on, että yksittäisessä potilastapauksessa kliiniseen päätökseen voi vaikuttaa kolme erilaista lääketieteellisen tiedon muotoa: kliininen tutkimusnäyttö, biologinen ja fysiologinen ymmärrys sekä kliininen kokemus. Tonelli myös esittää, että nämä kolme muotoa ovat perusluonteeltaan erilaisia, joten niitä ei voida asettaa arvojärjestykseen vaan niitä tulee yhdistää kulloisenkin tilanteen mukaan potilaan

parhaaksi. Jokaisella niistä on hänen mukaansa omat hyvät ja huonot puolensa, jotka vaikuttavat niiden sovellettavuuteen.³²

Kliinisten tutkimuksien vahvuutena on niiden toistettavuus ja kyky erotella yksittäisten interventioiden vaikuttavuuksia. Niillä on myös suurempi taipumus välttää tiettyjä harhoja. Huonona puolena taas on tulosten huono sovellettavuus yksittäiseen potilaaseen, kuten Tonelli toteaa: ”satunnaistetun kontrollitutkimuksen ei koskaan ole ollut tarkoitus vastata kysymyksiin tietyn [yksittäisen] potilaan hoidosta.”³²

Patofysiologisen päättelyn vahvuutena on sen kyky asettaa teoria intervention vaikutuksesta kriittiseen tarkasteluun. Esimerkiksi homeopatian tai tässäkin tutkielmassa käsitellyn retroaktiivisen rukouksen vaikutusten puolesta esitetyt tulokset voidaan turvallisesti tulkita sattumaksi, mikäli menetelmien puolustajilla ei ole esittänyt nykyistä uskottavampaa mallia vaikutuksen mekanismista. Biologinen päättely voi myös tukea kliinisessä tutkimuksessa löydetyn korrelaation herättämää ajatusta kausaliiteetista, jota pelkällä epidemiologisella tutkimuksella on vaikea osoittaa. Biologisen ja patofysiologisen päättelyn heikkous on kuitenkin siinä, että se kykenee selittämään asioita vain niin hyvin, kuin mitä sen hetkinen tieteellinen ymmärryksemme mahdollistaa. Tämän heikkouden havaitseminen on vaikeaa, sillä emme tiedä, mitä emme tiedä. Tämän vuoksi Tonelli toteaa, että nöyryys tämän luonteisen tiedon hyödyntämisessä on erittäin tärkeää.³²

Kliinisen kokemuksen merkittävin rooli taas on Tonellin mukaan diagnosoinnissa. Tämä on järkeenkäypää, koska keittokirjaa kirjaimellisesti seuratessa voidaan päätyä vain hyvin rajalliseen määrään ja vain nimenomaan tietynlaisia ruokalajeja. Lisäksi vasta käytännön kliinisessä työssä kyetään havaitsemaan, toimiiko tutkimusnäytön tukema hoito todella juuri tällä potilaalla. Myös klinikon anekdotaaliset havainnot hoitojen haitoista ja mielenkiintoisista poikkeamista yksittäisillä potilailla voivat toimia ponnahduslautana uuden tiedon syntymiselle. Heikkoutena on yksilöiden taipumus kognitiivisiin harhoihin.³² Lisäksi näkisin, että tämän tiedon subjektiivisuus johtaa siihen, että sen sovellettavuus riippuu täysin kyseessä olevan lääkärin ammattitaidosta ja valitettavan usein myös satunnaisista tekijöistä kuten vireystilasta jne.

Näkisin, että näiden kolmen tiedon lajin ominaispiirteiden, vahvuuksien ja heikkouksien ymmärtäminen on välttämätöntä niiden asianmukaiselle soveltamiselle. Niiden keskinäinen arvottaminen, tulkinta ja soveltaminen potilaan hoidossa on se, mikä viime kädessä erottaa lääkärin muista terveydenhuollon ammateista ja puoskareista.

7 Yhteenveto

Olen nyt tarkastellut näyttöön perustuvan lääketieteen käsitettä ja tehnyt johtopäätöksen, että sanana ”lääketiede” sisältää jo vaatimuksen ”näyttöön perustumisesta”, joten etuliite on merkityksetön samassa mielessä kuin ilmaus ”syötävä ruoka”. EBM:stä puhuttaessa tyypillisesti tarkoitetaan tutkimusten hierarkista luokittelua tutkimusmetodologian perusteella. Evidenssin käsite itsessään ei kuitenkaan tue tällaista ajattelua.

Olen tarkastellut lääketieteellisen tutkimuksen virhelähteitä ja osoittanut, että satunnaistaminen tutkimusmetodologisena menetelmänä soveltuu erinomaisesti indikaatioharhan välttämiseen. Tämä ei kuitenkaan tarkoita, että se on ainoa tapa välttää indikaatioharhaa, minkä lisäksi ajatus siitä, että satunnaistaminen automaattisesti tasaisi vertailuryhmät keskenään, on perusteeton. Satunnaistetun tutkimuksen arvottaminen yli muiden tutkimusmenetelmien voikin johtaa helposti tekniseen harhaan ja lääkkeellisten interventiodien perusteettomaan suosimiseen. Eri lääketieteelliset tutkimusmenetelmät onkin suunniteltu vastaamaan erilaisiin kysymyksenasetteluihin ja tutkimuksen taustalla tulisi aina olla uskottava ajatus vaikutusmekanismista.

Tämän vuoksi tärkeää olisi tarkastella lääketieteellisiä tutkimustuloksia kriittisesti ja kokonaisvaltaisesti ja ymmärtää kliinisen tutkimustiedon, patofysiologisesta päättelystä johdettavan tiedon sekä kliinisen kokemuksen laadullinen ero, jotta niitä voitaisiin käyttää asianmukaisesti lääketieteellistä tietoa arvioitaessa potilaan parhaaksi.

8 Johtopäätökset

Mitä tästä katsauksesta opimme? Mielestäni tärkein oppi on maltillisuus ja nöyryys johtopäätösten vetämisessä. Liian nopeat ja voimakkaat yleistykset ovat taipuvaisia johtamaan myös karkeampiin systemaattisiin sokeisiin pisteisiin. Sekä Bluhmin että Tonellin tarjoamissa lähestymistavoissa on havaittavissa kokonaisvaltainen tarkastelu: sekä ”perinteisessä” että ”näyttöön perustuvassa” lääketieteessä on vahvuutensa ja heikkoutensa, joiden tunnustaminen vaikuttaa luotettavimmalta tavalta edetä kohti totuutta tai ainakin sitä lähellä olevia vastauksia.

Tätä maltillisuutta ja ehdottomuuksien välttämistä ei kuitenkaan voi kutsua uudeksi, vaan enimmilläänkin unohdetuksi ajatukseksi, sillä jo satunnaistetun kokeen pioneeri Austin Bradford Hill totesi: ”Uskomus, että kontrolloitu koe on ainoa tapa [tuottaa lääketieteellistä tietoa], ei tarkoita, että heiluri olisi heilahtanut liian kauas ääriasentoon, vaan että se on kokonaan irronnut koukustaan.”³³

EBM-debattiin tutustuessani koin itse usein voimakkaan ristiriidan artikkeleissa käydyssä keskustelun ja käytännössä harjoitetun lääketieteen välillä. Vaikka evidenssihierarkian puolustajat ovat antaneet paljon myönnytyksiä eri metodologioiden tuottamalle tiedolle, on henkilökohtainen kokemukseni, että lääketieteen opetuksessakin kuin huomaamatta asetetaan satunnaistetut tutkimukset ja meta-analyysit epistemologisesti korkea-arvoisempaan asemaan perustelematta tätä valintaa sen enempää.

Tässä tutkielmassa on tietoisesti rajattu pois monia tärkeitä aiheita ja käsiteltyjenkin teemojen ääreltä löytyisi varmasti lisää tutkittavaa. Ehkä kiinnostavimpana ulkopuolelle rajattuna aiheena näen kysymyksenasettelun teknisestä harhasta ja tuottamamme tiedon luonteesta. Jos emme kykene käsittämään jotakin ilmiötä, se ei poista sen todellisuutta tai merkitystä. Kuinka siis saada luotettavaa tietoa lääketieteen piirissäkin vaikuttavista ilmiöistä, jotka eivät sovellu nykyisiin lääketieteen tutkimusasetelmiin? Ja laajemmin: kuinka konstruoida huonosti käsittämämme ilmiöstä luotettavaa tietoa antava tutkimusasetelma sen hetkisestä episteemisestä asemastamme käsin?

Samanaikaisesti täysin debatin kulusta riippumatta joka päivä lukemattomat potilaat tarvitsevat lääkäriltään linjauksen hoitoonsa. Näistä jokaisen yksilön tulisi saada, jos ei parasta, niin vähintään mielekkäintä tarjolla olevaa hoitoa vaivaansa ja näihin päätöksiin vaikuttaa ”evidenssin” lisäksi lukematon määrä muita käytännön asioita: vastaanottoajan kesto, potilaan oma vakaumus ja toive hoidosta, lääkärin mieliala ja suhtautuminen potilaaseen sekä tarjolla olevat hoitoresurssit ja poliittiset linjaukset siitä, mitä ja miten hoidetaan. “[Meidän olisi aika myöntää,] että lääketiede on hajanaista, poliittista sekä väistämättömästi normatiivista, ja vaikka oletamme evidenssin ottavan roolin päätöksenteossa, mitä todella tarvitsemme, on [moraalinen] päätös. (engl. judgement)”²¹

9 Lähdeluettelo

- ¹ Voelker R. Everything You Ever Wanted to Know About Evidence-based Medicine. JAMA 2015;313(18):1783.
- ² EBM working group. Evidence-Based Medicine A New Approach to Teaching the Practice of Medicine. JAMA 1992;268(17):2420.
- ³ Norman G. The paradox of evidence-based medicine. Commentary on Gupta (2003), A critical appraisal of evidence-based medicine: some ethical considerations. Journal of Evaluation in Clinical Practice 2003;9(2):129.
- ⁴ Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996.
- ⁵ Kvernbekk T. The concept of evidence in evidence-based practice. Educational theory 2011;61(5):515.
- ⁶ Achinstein P. The Book of Evidence. Oxford: Oxford University Press; 2001.
- ⁷ Borgerson K. Valuing Evidence: Bias and the Evidence Hierarchy of Evidence-Based Medicine. Perspectives in Biology and Medicine 2009;52(2):218.
- ⁸ Erwin E, Siegel H. Is confirmation differential? British Journal for the Philosophy of Science 1989;40(1):105
- ⁹ Tonelli M. The Philosophical Limits of Evidence-based medicine. Academic medicine 1998;73(12):1234.
- ¹⁰ Berg Friis JK, Rossel P, Slot Norup M (eds.). Philosophy of Medicine: 5 questions. Milton Keynes, Uk: Automatic Press 2011
- ¹¹ Lääkäriinvala - lääkäriliiton verkkosivut. 2013; Available at: <https://www.laakariliitto.fi/liitto/etiikka/laakarinvala/>. Accessed 4.12.2016, 2016.
- ¹² Shahar E. A Popperian perspective of the term 'evidence-based medicine'. Journal of Evaluation on Clinical Practice 1997;3(2):109.
- ¹³ Flaker G, Blackshear J, McBride R, Kronmal R, Halperin J, Hart R. Antiarrhythmic Drug Therapy and Cardiac Mortality in Atrial Fibrillation. JACC 1992;20(3):527.
- ¹⁴ Urbach P. The value of randomization and control in clinical trials. Statistics in medicine 1993;12:1421.
- ¹⁵ Fisher R. The Design of Experiments. 4th ed. Edinburgh: Oliver and Boyd; 1947.

-
- ¹⁶ Byar D, Simon R, Friedewald W, Schlesselman J, DeMets D, Ellenberg J, et al. Randomized Clinical Trials - Perspectives on Some Recent Ideas. *NEJM* 1976;295:74.
- ¹⁷ Bluhm R. From Hierarchy to Network: a richer view of evidence for evidence-based medicine. *Perspectives in Biology and Medicine* 2005;48(4):535.
- ¹⁸ GRADE working group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328.
- ¹⁹ Grossmann J, Mackenzie F. The Randomized Controlled Trial: gold standard, or merely standard? *Perspectives in Biology and Medicine* 2005;48(4):516.
- ²⁰ Vandembroucke J. Observational Research, Randomised Trials, and Two Views of Medical Science. *PLOS Medicine* 2008;5(3):0339.
- ²¹ Kerridge I. Ethics and EBM: acknowledging bias, accepting difference and embracing politics. *Journal of Evaluation in Clinical Practice* 2010:365.
- ²² Cho M, Bero L. The Quality of Drug Studies Published in Symposium Proceedings. *Annals of Internal Medicine* 1996;124(5):485.
- ²³ Gupta M. A critical appraisal of evidence-based medicine: some ethical considerations. *Journal of Evaluation in Clinical Practice* 2003;9(2):111.
- ²⁴ Leibovici L. Effects of remote, retroactive intercessory prayer on outcomes in patients with bloodstream infection: randomised controlled trial. *BMJ* 2001;323(1450).
- ²⁵ Worrall J. Evidence: Philosophy of science meets medicine. *Journal of Evaluation in Clinical Practice* 2010:356.
- ²⁶ Leibovici L. Author's reply. *BMJ* 2002;324:1037
- ²⁷ Giacomini M. Theory-Based Medicine and the Role of Evidence: Why the Emperor Needs New Clothes, Again. *Perspectives in Biology and Medicine* 2009;52(2):234.
- ²⁸ Merenstein D. Winners and Losers. *JAMA* 2008;291(1):15.
- ²⁹ Norman G. Examining the assumptions of evidence-based medicine. *Journal of Evaluation in Clinical Practice* 1999;5(2):139.
- ³⁰ Nisbett R, Ross L editors. *Human Inference: Strategies and Shortcomings in Social Judgement*. Prentice Hall; 1980.
- ³¹ Bluhm R. Some Observations on Observational research. *Perspectives in Biology and Medicine* 2009;52(2):252.
- ³² Tonelli M. The challenge of evidence in clinical medicine. *Journal of Evaluation in Clinical Practice* 2010:384.

³³ Hill AB. Reflections on the Controlled Trial. *Annals of the rheumatic diseases* 1965:107.