

H-indeks – et godt mål på forskningsaktivitet?

H-indeks brukes i økende grad som et mål på en persons publiserings- og forskningsaktivitet. H-indeks er beregnet ut fra antall publiserte artikler og det antall ganger hver artikkel er sitert. Det er viktig å kjenne til h-indeks og ikke minst hvilke begrensninger det er i å samle et helt forskerliv i kun ett tall.

Tor-Arne Hagve

tor.arne.hagve@ahus.no

Tverrfaglig laboratoriemedisin og medisinsk biokjemi

Divisjon for diagnostikk og teknologi

Akershus universitetssykehus

og

Institutt for klinisk medisin

Universitetet i Oslo

Jeg velger å innrømme det med en gang og uten blygsel; min er 18. Denne åpenheten skyldes ikke uttalt ekshibisjonisme, men heller at det her er snakk informasjon som helt og fullt er offentlig tilgjengelig. Alle som ønsker det kan med få tastetrykk få innsikt i både min, sin egen og andres størrelser. Når begrepet h-indeks blir bedre kjent i forskningsmiljøene, vil det sannsynligvis skape samme engasjerende behov for å kikke på andres aktiva som i den generelle befolkning når skattelister publiseres. Hovedhensikten med denne kronikken er å omtale h-indeks og diskutere begrensninger i bruken av denne.

H-indeks som mål på forskningskompetanse

H-indeks ble i 2005 foreslått av fysikeren Jorge Hirsch som et mål på en persons totale vitenskapelige aktivitet, herunder både produktivitet og impakt (1). H-indeks beskriver en persons vitenskapelige produksjon basert på totalt antall publikasjoner samt antall ganger hver artikkel er sitert. Definisjonen på h-indeks er: «en forsker har indeks h hvis h antall av forskerens totale antall publikasjoner har minst h antall siteringer og resten har færre enn h siteringer hver» (1). I klartekst, av mine 80 indekserte artikler har 18 artikler 18 eller flere siteringer og de resterende har mindre enn 18 siteringer.

Det har etter hvert blitt mye interesse for det ukompliserte i å forholde seg til bare ett tall, og de store publiseringsdatabasene som Web of Science (2) og Google Scholar (foreløpig kun til utprøving) (3) har nå etablert enkle måter for å få sitt og andres pass

påskrevet. Man velger database, skriver inn forskerens navn og frem kommer hele publikasjonslisten sortert etter antall siteringer med angivelse av blant annet h-indeks.

Jeg vil tro at leseren allerede nå har flere motforestillinger til h-indeks, både med tanke på troverdighet, nytte, verdi og ikke minst hvordan en slik størrelse kan misbrukes. Diskusjonen har selvsagt gått varmt i publiseringsverdenen, hvilket gjenspeiler seg i at antall siteringer av Hirsch's originale artikkel økte fra 50 i 2007 til over 650 i 2010 (4, 5).

De fleste studier av h-indeks er sammenlikninger med andre bibliometriske parametre, og stort sett finnes det som forventet god sammenheng mellom disse. Det er imidlertid også publisert en del studier om sammenhengen mellom h-indeks og resultat av klassisk fagfellevurdering. Bornmann & Daniel (6) fant at hos søkere på stipend på postdoknivå (n = 414) var h-indeks gjennomgående høyere for aksepterte søknader (gjennomsnitt 3,8) sammenliknet med refuserte (2, 7). Benway og medarbeidere undersøkte sammenhengen mellom h-indeks og de ulike akademiske nivåer innen urologi i USA og fant at «Professor» (n = 103), «Associate Professor» (n = 71) og «Assistant Professor» (n = 93) hadde h-indeks på henholdsvis (gjennomsnitt ± SD) 22,0 ± 10,7, 13,8 ± 6,9 og 8,0 ± 6,2 og forskjellene var statistisk signifikante (p < 0,0001) (7). Andre studier viser den samme trenden (8, 9).

Men en sak er å finne sammenheng mellom tidligere forskningsaktivitet og akademisk status, en helt annen sak er om en bibliometrisk størrelse kan si noe om sannsynligheten for fremtidig innsats. Og det er vel det egentlige målet både ved tildeling av forskningsmidler og ved ansettelse. Jorge Hirsch har foretatt en mye sitert undersøkelse av dette der han i en definert forskergruppe (n = 50) retrospektivt sammenliknet bibliometriske parametere i tre 12-årsperioder og fant at h-indeks har en betydelig større prediktiv styrke for en aktiv karriere enn de tidligere brukte bibliometriske parametrene (10). I andre studier er verdien av h-indeks ikke like tydelig (11, 12).

Andre bibliometriske mål for forskningskompetanse

Antall publikasjoner

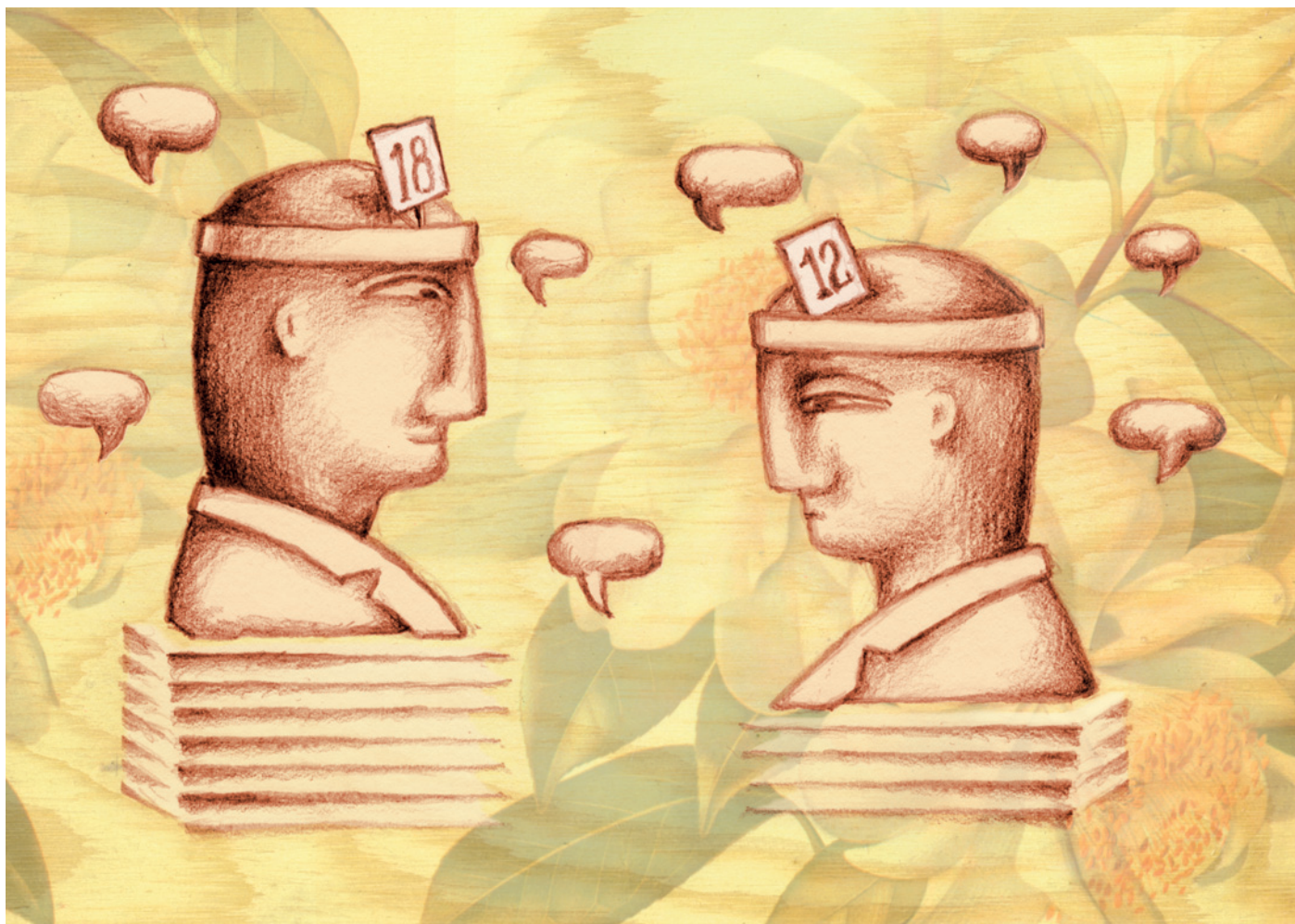
Innen mitt fagfelt og for 20 år siden var det enkle mål for en vellykket forsker å ha publisert 40 artikler i indekserte tidsskrifter før fylte 40 år, og 100 før fylte 50 år. Greit å forholde seg til, men selvsagt altfor enkelt og nærmest ubrukelig som eneste mål på forskningskompetanse. Hovedproblemet med antall publikasjoner som bibliometrisk størrelse er at den sier intet om kvalitet. Det er heller aldri dokumentert noen sammenheng mellom antall publikasjoner og forskningskvalitet (13, 14).

Totalt antall siteringer

Det antall ganger en persons artikler er blitt sitert, er heller intet godt mål på forskningsaktivitet, verken på produksjon eller på kvalitet. Et eksempel på det er H.O. Lowry som har publisert en av de mest siterte artikler noen gang (15). Han beskrev i 1951 en metode for bestemmelse av proteinkonsentrasjon, og dette arbeidet er sitert ca. 70 000 ganger (2). Lowrys metode viste seg å være anvendelig for kvantitering av protein i biologiske materialer, og uansett hva man ellers måler i slike materialer, blir resultatene veldig ofte relatert til proteinmengden. Lowrys metode ble standardmetode og er derfor ofte sitert. Et høyt antall siteringer kan altså være resultat av «one-hit wonders» og betyr ikke nødvendigvis god forskning over tid. Gjennomsnittlig antall siteringer per artikkel tar i noen grad også hensyn til produksjonen (antall artikler), men har samme svakheter som totalt antall siteringer. De ulike databaser rapporterer antall siteringer men ofte med stor variasjon mellom databasene. Antall siteringer er naturlig nok høyere for eldre artikler sammenliknet med nyere, er påvirket av selv-siteringer og varierer betydelig mellom fagområder og medisinske spesialiteter (13).

Impaktfaktor

Impaktfaktor er basert på gjennomsnittlig antall siteringer av hver enkelt artikkel publisert i et tidsskrift over en toårsperiode og er i utgangspunktet et mål på et tidsskrifts kvalitet. Selv om impaktfaktor ikke sier noe



Illustrasjon Stein Løken

om individuell forskningskompetanse, er den i mange år blitt brukt til vurdering av forskeres og institusjoners forskningsaktivitet, ikke minst pga. tilgjengelighet (Science Citation Index). Begrepet «high impact journals» er blitt en gjenganger i søknader, prosjektbeskrivelser og CV-er uten at det er åpenbart hvor grensen går mellom det beste, middelmådige og dårlige. Det må likevel sies at noen tidsskrifter skiller seg tydelig ut ved å ha en svært høy impaktfaktor, for eksempel har *New England Journal of Medicine* impaktfaktor 53,5. Til sammenlikning har de fleste skandinaviske (redigerte) internasjonale tidsskrifter en impaktfaktor i området 1–3. Det er også et faktum at nivået for impaktfaktor varierer betydelig mellom tidsskrifter innen ulike forskningsfelter og fagområder (16). Det er derfor ikke nødvendigvis riktig eller nyttig å benytte impaktfaktor ved sammenlikning av forskningsaktivitet innen ulike fagområder. Mange av de samme usikkerhetsfaktorene som er knyttet til h-indeks, gjelder også for impaktfaktor (16).

Innvendinger mot h-indeks

Er så h-indeks et bedre bibliometrisk mål for forskningsaktivitet enn de tidligere stør-

relsene? Mange mener ja, fordi den bedretar hensyn til både produksjon og kvalitet. Men som sagt, det er mange motforestillinger og mye usikkerhet i fagmiljøene. I det følgende beskrives de viktigste begrensningene for bruk av h-indeks.

Selvsitering

Selvsitering er rapportert å være i størrelsesorden 6–15% av alle siteringer innen medisinsk publisering, med store forskjeller mellom fagområder og tidsskrifter (17–19). Selvsitering kan altså representere en temmelig stor andel av siteringer og innvirkningene på bibliometriske størrelser, inkludert h-indeks er åpenbare, ikke minst som basis for ulikheter mellom fagområder.

I hvor stor grad kan jeg taktisk og aktivt øke min h-indeks ved selvsitering? Jeg har altså 18 artikler som har flere enn 18 siteringer, og hvis jeg i mine neste to artikler siterer den som har 17 siteringer, vil jeg ha h-indeks på 19. Hvis jeg siterer den enda en gang, kan min h-indeks raskt bli 20. Her kan det altså brukes taktikk, og for sin karriere er man kanskje villig til det? Men på sikt gir denne type triksing liten uttelling for en aktiv forsker, fordi de naturlige siteringer som kommer ved en jevn publisering

av nye artikler, snart vil ta igjen de taktiske siteringene (17, 18). Det hevdes også at bare 10% av alle h-indeksøkende enkelt-siteringer har betydning for h-indeks etter 9 md. Den taktiske siteringen må eventuelt aktivt følges opp, og det lar seg sannsynligvis ikke gjøre i lengden.

Førsteforfatter versus medforfatter

Innen fagområdet medisinsk biokjemi er det vist at bare 20–30% av alt som publiseres, er resultater fra egeninitierte prosjekter (20). Dette fremkommer etter gjennomgang av publikasjonslister fra de fleste universitetssykehus og en del andre store sykehus i Norden. For de fleste artikler som utgår fra fagområdet medisinsk biokjemi er det altså snakk om medforfatterskap, der graden av vitenskapelig involvering ikke nødvendigvis er åpenbar eller dokumentert. Det er ingen bibliometriske størrelser, inkludert h-indeks som tar hensyn til at det sannsynligvis er førsteforfatter (ofte i samarbeid med sisteforfatter) som har bidratt med den faglige bakgrunnen for prosjektet, skrevet protokollen, gjennomført datainnsamling og skrevet artikkelen. Dette i motsetning til en medforfatter som i enkleste fall bare har bidratt med et datasett for en

analyse. For å øke min h-indeks kan jeg etablere en analyse som bare få andre utfører og markedsføre den til aktive forskningsgrupper, med medforfatterskap som gevinst.

Antall forfattere

Å være en av mange forfattere på en artikkel kan indikere gode samarbeidsevner og stor vitenskapelig kontaktflate. Men hver forfatters innsats er da nødvendigvis mindre enn om det hadde vært færre forfattere. Til tross for at Hirsch tar opp dette som et problem i sin opprinnelige artikkel fra 2005 (1), er antall forfattere ikke tatt hensyn til i beregning av h-indeks eller i noen av de videreutviklede indeksene (21). Et stort antall forfattere kan også ha positiv betydning for de respektive forfatters h-indeks på lengre sikt fordi det da er flere som kan selvsitere den samme artikkelen.

Alder

Det er åpenbart at h-indeks favoriserer eldre forskere både fordi de har publisert flere artikler og fordi det har vært mulig å selvsitere en større andel av artiklene over lengre tid. H-indeks sier heller intet om tidsaspektet for forskningsaktivitet. I den sammenhengen er det viktig å være bevisst at h-indeks aldri kan bli mindre. Slike forhold kan spesielt ha betydning ved vurdering av søkere til forskerstillinger. Det er foreslått å kompensere for ulik alder ved å dividere h-indeks med antall forskerår (hvilket gir såkalt m-kvotient), bestemt fra året den første artikkelen ble publisert (21). M-kvotient kan enkelt beregnes, men brukes lite.

Forskjell mellom fagområder

Det er betydelige forskjeller mellom fagområder og innen fagområder når det gjelder publiseringsaktivitet (antall artikler per år), omfanget av hver publisering og antall siteringer per artikkel. Det generelle bildet er at det innen realfagene, inkludert medisin, publiseres flere artikler, og hver artikkel er kortere og siteres hyppigere enn artikler fra humanistiske fag (22). Men det er også store forskjeller innen realfagene. Gjennomsnittlig antall siteringer per artikkel i årene 1995–2005 (fra Institute for Scientific Information, ISI) var 24,6 for molekylær biologi, 19,6 for immunologi, 14,0 for mikrobiologi, 10,6 for klinisk medisin og bare 2,7 for matematikk (23). Den åpenbare konsekvensen av dette er at det er vanskelig å sammenlikne h-indeks mellom ulike fagområder, og det er endog store forskjeller mellom ISI-definerte grupper (medisinske spesialiteter) relatert til medisin.

Originalartikkel versus oversiktsartikkel

Ulike artikkeltyper som originale artikler, oversiktsartikler og kasuistikker teller like mye ved beregning av h-indeks. Ved

å satse på oversiktsartikler kan jeg øke min h-indeks i en fart, særlig hvis jeg skriver om et hett tema. Da slipper jeg å bruke tid og ressurser til langvarig innsamling av pasientdata, teknisk krevende analyser og avansert statistisk behandling samt å skrive søknader om finansiering og søke godkjenning fra regional etisk komite. Det er vel bare å lese noen artikler om et rimelig nytt emne, for eksempel om nye bibliometriske aktivitetsindikatorer innen forskning?

Kontekst

H-indeks tar ikke hensyn til i hvilken sammenheng en artikkel er sitert. Et eksempel på at dette er et problem, er at en kontroversiell artikkel som ikke nødvendigvis er av høy vitenskapelig kvalitet eller har faglig relevans, kan siteres ofte i forsøk på å belyse en sak fra mange vinkler, eller for å sprite opp ens egen innledning. Likeledes tas det ikke hensyn til siteringer av artikler som er trukket tilbake eller hvor det er påvist forskningsfusks. Problemet med kontekst gjelder også for andre bibliometriske størrelser.

Ulike databaser gir ulik h-indeks

Det er vist at de ulike siteringsdatabaser rapporterer ulik h-indeks for samme forsker (4, 24). Som nevnt har jeg ifølge Google Scholar publisert 80 artikler med h-indeks 18, mens jeg i Web of Science er registrert med 71 artikler og h-indeks 20. Dette skyldes delvis at siteringene registreres på forskjellige tidspunkter og delvis at databasene har forskjellige kriterier for hvilke tidsskrifter og ikke minst hvilke kongressabstrakter som registreres. Det viktige er at man bruker h-indeks fra samme database ved sammenlikning av publiseringsvirksomhet.

Oppsummering

Kan h-indeks anbefales som et troverdig og nyttig mål på forskningsaktivitet? Uansett hva man måtte mene om det, brukes h-indeks allerede som mål på forskeres aktivitet, blant annet ved vurdering og rangering av søkere til akademiske stillinger. Det er derfor viktig at man både innen og utenfor forskningsmiljøene har kunnskaper om og forståelse for alle de begrensninger som alltid vil være knyttet til denne type forenklinger. Hvis h-indeks skal brukes som mål på forskningsaktivitet, må den ses i sammenheng med andre bibliometriske parametere, og være et supplement til fagfellevurdering.

Tor-Arne Hagve (f. 1953)

er overlege i medisinsk biokjemi ved Divisjon for diagnostikk og teknologi og professor II ved Institutt for klinisk medisin.

Ingen oppgitte interessekonflikter.

Litteratur

- Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A* 2005; 102: 16569–72.
- Web of knowledge. www.webofknowledge.com [27.10.2011].
- scHolar index. <http://interaction.lille.inria.fr/~rousseau/projects/scholarindex/index.cgi?btnG=Search+> [27.10.2011].
- Bornmann L, Daniel HD. The state of h index research. Is the h index the ideal way to measure research performance? *EMBO Rep* 2009; 10: 2–6.
- Bornmann L, Marx W. The h-index as a research performance indicator. *European Science Editing* 2011; 37: 77–80.
- Bornmann L, Daniel HD. Does the h-index for ranking of scientists really work? *Scientometrics* 2005; 65: 391–2.
- Benway BM, Kalidas P, Cabello JM et al. Does citation analysis reveal association between h-index and academic rank in urology? *Urology* 2009; 74: 30–3.
- van Raan AF. Comparison of the Hirsch-index with standard bibliometric indicators and with peer judgement for 147 chemistry research groups. *Scientometrics* 2006; 67: 491–502.
- Lovegrove BG, Johnson SD. Assessment of research performance in biology: How well do peer review and bibliometry correlate? *Bioscience* 2008; 58: 160–4.
- Hirsch JE. Does the H index have predictive power? *Proc Natl Acad Sci U S A* 2007; 104: 19193–8.
- Poynard T, Thabut D, Munteanu M et al. Hirsch index and truth survival in clinical research. *PLoS ONE* 2010; 5: e12044.
- Hönekopp J, Kleber J. Sometimes the impact factor outshines the H index. *Retrovirology* 2008; 5: 88.
- Arencibia-Jorge R, Barrios-Almaguer I, Fernandez-Hernandez S et al. Applying successive h-indices in the institutional evaluation: a case study. *Am J Soc InfoSci Technol* 2008; 59: 155–7.
- Patel VM, Ashrafiyan H, Ahmed K et al. How has healthcare research performance been assessed?: a systematic review. *J R Soc Med* 2011; 104: 251–61.
- Lowry OH, Rosebrough NJ, Farr AL et al. Protein measurement with the Folin phenol reagent. *J Biol Chem* 1951; 193: 265–75.
- Seglen PO. Why the impact factor of journals should not be used for evaluating research. *BMJ* 1997; 314: 498–502.
- Bartneck C, Kokkelmans S. Detecting h-index manipulation through self-citation analysis. *Scientometrics* 2011; 87: 85–98.
- Kulkarni AV, Aziz B, Shams I et al. Author self-citation in the general medicine literature. *PLoS One* 2011; 6: e20885.
- Engqvist L, Frommen JG. The h-index and self-citations. *Trends Ecol Evol* 2008; 23: 250–2.
- Hagve TA. Publiseringsmønsteret innen fagområdet medisinsk biokjemi i Norden. *Klinisk Biokjemi i Norden* 2005; 3: 32–34.
- Alonso S, Cabrerizo FJ, Herrera-Viedma E et al. H-index: a review focused on its variants. Computation and standardization for different scientific fields. *J Informatics* 2009; 3: 273–89.
- Harzing AW. Reflections on the h-index. www.harzing.com/pop_hindex.htm [23.9.2011].
- Iglesias JE, Pecharramán C. Scaling the h-index for different scientific ISI fields. *arXiv: physics/0607224* [27.10.2011].
- Thompson DF, Callen EC, Nahata MC. New indices in scholarship assessment. *Am J Pharm Educ* 2009; 73: 111.

Mottatt 11.7. 2011, første revisjon innsendt 26.9. 2011, godkjent 20.10. 2011. Medisinsk redaktør Erlend T. Aasheim.