

Rosława Revheim
Główna Biblioteka Lekarska, Warszawa

Impact Factor – polubić czy lekceważyć?

Streszczenie

W celu identyfikacji wartościowych publikacji naukowych i medycznych stosuje się metody bibliometryczne, jak: obliczanie wskaźników cytowań, impact factorów, zakresu obiegu czasopism i wiele innych parametrów o charakterze statystycznym. Metody te - równoległe z ich rozwojem - budzą liczne kontrowersje w świecie nauki jako potencjalnie stroniczne, łatwo ulegające manipulacji lub wręcz niosące zafałszowane informacje

Mimo to impact factor i wskaźniki pochodne są powszechnie stosowane jako mierniki jakości naukowej produkcji piśmienniczej i nic nie wskazuje na to, by w najbliższym czasie miało się tu coś zmienić.

W Głównej Bibliotece Lekarskiej na zlecenie Departamentu Nauki i Kształcenia MZiOS opracowano analizę bibliometryczną piśmiennictwa medycznego polskiego za lata 1988-1998 w postaci corocznych raportów zawierających wyniki analiz materiałów nadsyłanych przez biblioteki placówek resortowych, tj. akademii medycznych, CMKP i instytutów MZiOS.

Udział rezultatów analizy bibliometrycznej w procesie decyzyjnym rozdziału środków na programy badawcze wynosił ok. 25%, co jest najlepszym dowodem na celowość pracy nad uzyskiwaniem możliwie obiektywnych wyników takich analiz.

Niezwykle trudnym, czy wręcz niemożliwym zadaniem dla ludzi nauki – w przypadku medycyny: klinicystów, badaczy, analityków polityki zdrowotnej – jest śledzenie na bieżąco ponad 2 mln publikacji naukowych ukazujących się każdego roku w czasopiśmie medycznych i naukowych.

Ilościowe pomiary produkcji piśmiennictwa naukowego są powszechną i najłatwiejszą metodą oceny stanu nauki w poszczególnych dziedzinach, jednakże liczba publikacji sama w sobie nie informuje o jakości pracy naukowej. Informacje o jakości niosą dwie składowe procesy naukowego: zamieszczenie publikacji w czasopiśmie o wysokim prestiżu w świecie nauki, oraz jej cytowanie przez innych autorów, co oznacza przydatność pracy naukowej, lub – innymi słowy – jej wpływ na innych autorów mierzony ilością cytowań.

Te fakty leżą u podstaw bibliometrycznych metod oceniania stanu nauki poprzez indeksy cytowań oraz impact factory czasopism.

O ile jednak intencją bibliometrii było zapewnienie obiektywnej metody oceny stanu nauki, o tyle znaczenie indeksów cytowań i impact factorów czasopism nie zawsze znajduje zrozumienie, a reakcje na nie wahają się od wyraźnego przeceniania ich roli do sceptycznego niedowierzania.

Z drugiej strony doświadczenie wykazuje, że w każdej dziedzinie nauki najlepsze czasopisma to te, w których droga do druku jest najtrudniejsza i one też mają wysoki impact factor. Czasopisma te powstały na długo przed tym nim impact factor został wymyślony, a jego stosowanie jako miary jakości pracy naukowej jest dlatego powszechne, że dobrze współgra z opinią środowiska oceniającą które czasopisma są naukowo najlepsze w danej dziedzinie.

Impact factor opiera się na informacji otrzymywanej z indeksów cytowań. Termin „impact factor” jest związany z opublikowaniem w 1963 roku Indeksu Cytowań Naukowych (Science Citation Index) za rok 1961. Od tego czasu rośnie baza danych zawierająca cytowania, które pojawiają się w bibliografiach artykułów w ogromnej liczbie czasopism naukowych. Baza ta jest produkowana przez Instytut Informacji Naukowej (Institute for Scientific Information) w Filadelfii, USA.

Impact factor (IF) dla czasopisma jest obliczany przez podzielenie liczby cytowań artykułów z tego czasopisma przez liczbę artykułów opublikowanych przez to czasopismo w tym samym okresie 2 lat. Inaczej mówiąc – IF czasopisma jest liczbą cytowań artykułów cytowanych podzieloną przez liczbę artykułów, które mogłyby być cytowane.

Impact factory czasopism są publikowane w postaci listy za dany rok w Science Citation Index – Journal Citation Reports, a końcowym tego skutkiem jest szerokie stosowanie IF do porównywania jakości czasopism oraz publikacji w nich zamieszczonych.

Impact factor stał się przedmiotem licznych kontrowersji, gdy stopniowo przekształcił się w kierunku oceny zarówno czasopism jak i autorów publikacji, a także szacowania i porównywania poziomu pracy naukowców i instytucji naukowych.. Tymczasem proste porównywanie IF czasopism może być zwodnicze, jak choćby w przypadku czasopism o charakterze przeglądowym, które często mają wyższy IF od tych, które publikują prace oryginalne.

Dokładniejsza analiza zestawu czasopism biomedycznych wyselekcjonowanych do Journal Citation Reports ukazuje kilka charakterystycznych cech: na skali IF wyżej znajdują się czasopisma stricte naukowe niż kliniczne; anglojęzyczne niż publikowane w innych językach; amerykańskie niż europejskie; oraz: najbardziej prestiżowe czasopisma z różnych dziedzin czy specjalności mają bardzo różne IF ;ponadto wolny dostęp online do czasopisma podnosi jego impact factor.

Zarówno zwolennicy jak i przeciwnicy stosowania IF przyznają zgodnie, że istnieje wiele ograniczeń tego systemu, które trzeba uwzględnić we wszystkich analizach dokonywanych z jego użyciem. I tak:

1. IF odzwierciedla poziom czasopisma raczej niż artykułów w nim zamieszczonych. Przyjęcie założenia, że wszystkie artykuły w danym czasopiśmie są podobnej jakości jest oczywiście błędne. Nikt przecież nie porównuje jakościowo pracy przeglądowej z pracą oryginalną wnoszącą nowe treści do nauki, a zatem przyczyniającą się do jej postępu.
2. Zmiany w kierunkach zainteresowań w nauce wpływają na IF czasopism. Dotyczy to zwłaszcza dziedzin rozwijających się mniej dynamicznie niż np. biologia molekularna, czy genetyka, ale przecież nie „gorszych” czy też mniej ważnych dla nauki w ogóle. Przy tym metoda obliczania IF dyskryminuje autorów odwołujących się do prac opublikowanych wcześniej niż w okresie poprzednich dwóch lat.
3. IF nie mówi nic o rzetelności procesu doboru artykułów do publikacji. Wydawcy – częściowo odpowiedzialni za IF czasopisma –mogą brać pod uwagę przewidywalny wskaźnik cytowań danego artykułu przy podejmowaniu decyzji o jego opublikowaniu lub nie.
4. Ze względu na strukturę bazy SCI, która nie rozróżnia artykułów oryginalnych, prac przeglądowych, abstraktów z konferencji czy też listów do redakcji, IF może być względnie łatwym obiektem manipulacji zarówno przez autorów jak i wydawców; np. zawyżanie IF przez samocytowania lub zalecanie autorom cytowanie artykułów zamieszczonych w danym czasopiśmie.

Generalnie – wielkość IF określa przybliżony status naukowy czasopisma czyli „Nature” z IF 25,8 jest postrzegane jako lepsze naukowo czasopismo, niż np. „Gastroenterology” z IF 12,2.

Jednakże IF nie daje całościowej oceny jakości czasopisma. Znakomitym potwierdzeniem tej tezy jest badanie ankietowe przeprowadzone w Narodowych Instytutach Zdrowia w USA, gdzie poproszono 50 naukowców o sporządzenie listy

najbardziej prestiżowych czasopism naukowych i klinicznych. Na liście rankingowej znalazły się czasopisma, których IF wahały się od 3 do 168 ułożone w porządku zupełnie niezależny; od wartości ich impact factorów.

Powstaje tu nieco drażliwe pytanie – kto jest uprawniony do stwierdzenia, która z metod oceny jakości czasopisma jest lepsza – system IF czy lista rankingowa ułożona przez autorytety naukowe?

Poza tym proste porównywanie IF fałszuje obraz, o ile dwa porównywane czasopisma nie konkurują o publikowanie tych samych artykułów, lub porównuje się czasopisma z różnych specjalności medycznych.

A jednak wymienione ograniczenia, czy też wady, jak się zdaje, nie pomniejszają wartości IF jako miary.

Impact factor został zaprojektowany w celu obiektywnego porównywania poziomu czasopism i – uwzględniając ograniczenia spełnia tę rolę dobrze. Reprezentuje jakościowy poziom czasopism i, jednocześnie, jest wygodnym w użyciu parametrem przy wyborze odpowiedniego czasopisma do ogłoszenia wyników pracy naukowej, której integralną częścią jest oczywiście publikacja.

Na uniwersytecie w Atenach przeprowadzono niedawno badanie wśród 438 anestezjologów z Kanady, USA i Europy dotyczące wartości i ważności IF jako miernika jakości czasopism. Okazało się, że wprawdzie europejscy lekarze nieco wyżej oceniają ważność IF, jednak wszyscy zgadzają się co do tego, że publikowanie artykułów w czasopismach o wysokim IF ułatwia zdobywanie funduszy na badania.

Istnieją co najmniej dwa powody, dla których IF jest i będzie jeszcze długo używany: po pierwsze – jest najlepszym, najbardziej obiektywnym z dostępnych miernikiem; po drugie – przykładanie wagi do jakości czasopisma zmusza autorów do koncentrowania się raczej na wysokiej jakości publikacji niż dużej ich ilości.

Ponadto statystycznie udowodniono, że mierniki aktywności naukowo-piśmienniczej są od siebie liniowo zależne, a zatem wystarczy brać pod uwagę jeden możliwie kumulatywny wskaźnik udziału w międzynarodowym obiegu informacji, jakim jest np. sumaryczny IF.

Podsumowując: system impact factora jest obciążony pewną niespójnością logistyczną i nie może być traktowany jako jedyne kryterium oceny aktywności naukowej autorów publikacji i ważności czasopism. Z drugiej strony żaden lepszy system nie został dotąd opracowany i wdrożony do praktyki bibliometrycznej. A zatem, jak wszystkie mierniki statystycznej natury, IF winien być stosowany rozważnie, ze zrozumieniem mocnych i słabych stron systemu oraz ze zdrowym rozsądkiem przy interpretacji wyników analiz bibliometrycznych.

Główna Biblioteka Lekarska ma kilkuletnie własne doświadczenia w zakresie analiz bibliometrycznych piśmiennictwa medycznego polskiego z użyciem impact factorów i indeksów cytowań.

Na początku 1991 roku w Departamencie Nauki i Kształcenia Ministerstwa Zdrowia i Opieki Społecznej, którego dyrektorem był w tym czasie dr Rafał Niżankowski, powstała inicjatywa opracowania systemu wartościowania (rankingu) pracowników nauk i instytucji medycznych – na tyle obiektywnego, by mógł służyć procesowi decyzyjnemu przy podziale środków pieniężnych przyznawanych przez MZiOS na programy badawcze.

Realizacja tego zadania została zlecona GBL, gdzie w marcu 1991 został opracowany Projekt Systemu Informacji o publikacjach Pracowników Naukowych Resortu Zdrowia. Osobą odpowiedzialną za wdrożenie projektu był ówczesny kierownik

Zakładu Naukowej Informacji Medycznej dr Dariusz Kuźmiński.

Analizę bibliometryczną polskiego piśmiennictwa medycznego rozpoczęto od roku 1988. Narzędziami pracy w programie analiz były przede wszystkim:

- Indeks Cytowań (Science Citation Index on CD-ROM), który GBL posiada za lata 1988-95 plus I kwartał 1996;
- Lista Rankingowa Czasopism (Key Figures from the Journal Rankings – SCI Journal Citation Reports) w postaci drukowanej.

Materiałem analizowanym były dane nadsyłane z placówek resortowych, tj. akademii medycznych, CMKP i instytutów MZiOS, a dotyczące średnio rocznie 2500 publikacji w czasopismach z impact factorem większym od zera.

Rezultaty analiz przedstawiano w postaci arkuszy kalkulacyjnych i wykazów autorów o największej w danym roku liczbie cytowań, najczęściej cytowanych artykułów, artykułów opublikowanych w najbardziej prestiżowych czasopismach, autorów o największym skumulowanym (sumarycznym) impact factorze, etc.

W ocenie Departamentu Nauki MZiOS udział wyników analiz bibliometrycznych w procesie decyzyjnym rozdziału środków na programy badawcze wynosił ok. 25%, co jest najlepszym potwierdzeniem celowości podejmowania wysiłku uzyskiwania możliwie obiektywnych wyników prac w zakresie bibliometrii – przy całej jej złożoności i wielu kłopotach z pozyskaniem pełnych danych do analizy.

Stan obecny bibliometrii w Głównej Bibliotece Lekarskiej prezentuje się bardzo mizernie.

Od 1996 roku ze względu na brak funduszy GBL nie zakupuje bazy SCI on CD-ROM.

Kłopoty finansowe spowodowały także brak modernizacji sprzętu komputerowego, który w większości pochodzi sprzed 10 lat i więcej.

Jednocześnie MZ generalnie i także z powodów finansowych, straciło zainteresowanie programem bibliometrycznym, a większość bibliotek medycznych traktuje przygotowywanie materiałów do analiz jako dodatkowe i mało przydatne obciążenie.

Rewitalizacja programu analiz bibliometrycznych w Głównej Bibliotece Lekarskiej kosztowałaby małą fortunę, której nie ma ani ministerstwo ani sama GBL, chętnych do sponsoringu też nie widać. W ten sposób jeszcze jedna wartościowa inicjatywa rozbiła się o mur ekonomicznej niewydolności systemu rozwoju nauki w Polsce.

Warszawa, 2002 r.

Bibliografia

1. Callahan M.: Journal prestige, publication bias and other characteristics associated with citation of published studies in peer-reviewed journals. JAMA, June 5, 2002, vol. 287, No 21.
2. Lee K. P.: Association of Journal Quality Indicators with Methodological Quality of Clinical Research Articles. JAMA, June 5, 2002, vol. 287, No 21.
3. Whitehouse C. H.: Citation rates and impact factor: should they matter? The British Journal of Radiology, Jan 2001, vol. 74
4. Fassoulaki A.: Academic anesthesiologists' views on the importance of impact factor of scientific journals: a North American and European survey. Canadian Journal of Anaesthesiology, Nov 2001, 48 (10).