

Jest nie najlepiej, a może być bardzo źle.

Pozwolę sobie w tym tekście zwrócić uwagę na niebezpieczeństwa systemu tzw. oceny parametrycznej i częściowo odnieść się do pewnych tez tekstu prof. Andrzeja Białynickiego-Biruli.

1. Czy ocena musi być parametryczna? Otóż nie, czego przykładem jest opisany w artykule prof. Pacha system brytyjski (por. Forum Akademickie <https://forumakademickie.pl/fa/2014/10/pl-vs-uk/>). System brytyjski, ekspercki w swej istocie, wymaga zapewne wiele pracy, ale przecież ocena jednostek nie musi być częsta – nauka to działania długodystansowe, więc ocena raz na 6-8 lat byłaby wystarczająca. System brytyjski w większym stopniu ocenia osiągnięcia – to co jest istotą nauki, a nie pośrednie wskaźniki jak publikacje, cytowania itp. W systemie polskim wkrótce pytanie: no dobrze, ale co wynikło (w sensie wyników naukowych) z działalności danej jednostki będzie zbędne. Nie oznacza to, że proponuję przeniesienie dokładne tego systemu ani, że realne jest dokonanie kategoryzacji w 2017 w jakiejś wersji tego systemu. Chcę jednak podważyć aksjomat, że system musi być „parametryczny”.
2. Czy nie ma innych propozycji na stole? Słyszałem kilka razy propozycję prof. Błockiego: patrzmy na liczbę, ewentualnie wielkość uzyskanych grantów zewnętrznych poważnych instytucji – ponieważ w każdym postępowaniu grantowym jest ekspercka ocena dotychczasowych osiągnięć, więc uzyskane wyniki dają sumaryczną ocenę poziomu jednostki (też pośrednią). Niestety ze względu na ograniczanie finansowania przez NCN ten system byłby bardzo ostry dla wielu jednostek, ale mógłby przyzwoicie działać, gdyby procent sukcesu w NCN wynosił więcej. Warianty tego systemu mogłyby być użyteczne. Wyżej wspomniany system brytyjski w jakiejś odmianie też pewnie dobrze by służył polskiej nauce: dlaczego nie możemy wzorować się na najlepszych? Możliwość jest zapewne więcej.
3. Czy system oceny pośredniej (taki jak obecnie) uznający, że poziom naukowy jest dobry skoro powstało wiele publikacji w czasopiśmie jest właściwy? Otóż *Wielkie Uproszczenie* wg. prof. Białynickiego-Biruli jest niestety w niektórych naukach zabójcze, a w innych obawiam się, że będzie. Byłem świadkiem dyskusji na posiedzeniu Komitetu Matematyki z członkiem komisji czasopism (to ci co wypracowali system punktów) – mimo dość dramatycznych wypowiedzi prof. Łuczaka o sytuacji informatyki (też nauka matematyczna), gdzie jego zdaniem rola czasopism jest drugorzędna, a liczą się publikacje na ważnych konferencjach, nie było nawet próby zrozumienia problemu. To jest problem, bo stawia przed autorem pracy wybór – czy ma opublikować tam gdzie to będzie naprawdę zauważone, czy tam gdzie ucieszy to Dziekana i da profity jednostce. Taki sam lub analogiczny problem dotyczy, o ile wiem, także wielu nauk technicznych. W naukach matematycznych system taki jak jest, będzie zdecydowanie nadmiernie promował grupy, które będą publikowały w słabych czasopiśmie o wysokim IF. Celem staje się IF a nie nauka.
4. Czy w ogóle wskaźniki bibliometryczne są dobre do oceny: tu powołałam się na opinię prof. Jackowskiego: *Nie znajduje empirycznego potwierdzenia rozpowszechniony pogląd, że proste wskaźniki bibliometryczne są „dobre” do oceny większych jednostek naukowych. Wiadomo od dawna, że tak nie jest nawet w makroskali.* Prof. Jackowski powołuję się tu na bardzo głęboką analizę o wiele mówiącym tytule: P. O Seglen,

“Why the impact factor of journals should not be used for evaluating research”, British Journal of Medicine 314 (1997), 498-502.

5. Wreszcie dochodzimy do pytania: Czy system musi być jednakowy dla wszystkich nauk? Moim zdaniem nie może być jednakowy, jeśli jest pośredni – mogą być wspólne ramy (np. ustalone kategorie, co one oznaczają w porównaniu do nauki światowej) ale nie jednakowe kryteria – jak czytam w systemie brytyjskim taka elastyczność jest przewidziana. Wskazane by było przynajmniej, aby eksperci przygotowali konkretne kryteria dla poszczególnych dziedzin nauki.
6. Czy system bardziej ekspercki doprowadzi do manipulacji i oszustw? Tak pewnie może się stać ale i obecny system już doprowadza do manipulacji. Publikuje się mnóstwo, niektórzy szukają słabych czasopism o wysokim IF itp. Każdy system jest nieodporny na nieuczciwość, ale nie ma innego wyjścia jak walczenie z tym, piętnowanie patologii itd. Problemem jest też odpowiednie przygotowanie ekspertów do pełnienia takiej roli, ustalenie ram ich funkcjonowania – doświadczenia NCN, a wcześniej KBN mogłyby być tu bardzo przydatne.
7. Czy obecny system daje jasny przekaz: *wyniki prac naukowych należy publikować w czasopismach o najwyższej randze*? Obawiam się, że przekaz jest inny: należy publikować wiele (rozdrabnianie publikacji) najlepiej w nuncie gdzie jest wiele publikacji, cały przemysł publikacji (to zapewnia cytowania) i w czasopismach o wysokim IF (która to kategoria zawiera czasopisma wysokiej rangi, ale wiele marnych). To już brzmi gorzej, nieprawdaż? Jaki dajemy przekaz naszym młodszym koleżankom i kolegom? I ku czemu to wszystko zmierza?
8. Czy wyniki oceny (dotychczasowych ocen parametrycznych) nie budzą zasadniczych protestów? Wielu z Państwa wie, że przy ostatniej kategoryzacji istniała realna szansa, że pewna silna matematycznie jednostka, mająca wiele osiągnięć również w matematyce uzyskałaby zapewne w GWO matematyka zdecydowanie lepszy wynik niż wszystkie oceniane w tej GWO jednostki. Jednostka ta, podkreślam, jest matematycznie bardzo dobra, ale zapewne nadzwyczajny jak na matematykę wynik zawdzięczałaby publikacjom z innych nauk, gdzie jest inny wzorzec cytowania (tak przynajmniej sędzę na podstawie publicznie dostępnych danych). Ostatecznie jednostka ta była oceniana w innej GWO i została tam oceniona wysoko ale nie najwyżej. Inna decyzja (tj. pozostawienie tej jednostki w GWO matematyka) była z pewnością w systemie do pomyślenia. Czy wtedy też środowisko uznałoby, że wynik jest prawidłowy? Ile mniej spektakularnych przykładów tego typu można by znaleźć? Sprawa podziału na GWO jest dość kluczowa i może przynieść łatwo niespodzianki. Jednym z problemów jest to, że wynik punktowy przyzwoity w jednej nauce może dać znakomity wynik w innej nauce (bez związku z poziomem uprawiania tych nauk), właśnie ze względu na różne kultury publikowania i cytowania w różnych dziedzinach nauki (por. artykuł prof. Błockiego i Życzkowskiego w Nauka 2/2013).

Gdy otrzymamy zaskakujące wyniki kolejnej oceny będzie za późno aby protestować – odpowiedź będzie brutalna: uczciwy i wspaniały system dał taki wynik, a feudalne środowiska naukowe nie chcą dostrzec, że dotychczasowe hierarchie były fałszywe, oparte na subiektywnych kryteriach jeśli nie wręcz na nieuczciwych powiązaniach.

Co należałoby zmienić? – obawiam się, że wszystko, ale proste rozwiązania jak zwykle są mało realne czy wręcz nie realne – nie widać sojuszników a i środowisko matematyczne też wykazuje ograniczoną aktywność w tej sprawie. Pobawmy się więc w mały realizm:

1. Zastąpić IF innym parametrem: trudniejszy w manipulacji jest Article Influence Score. Doskonała analiza zalet tego wyboru jest w artykule: Z. Błocki, K. Życzkowski, „Czy można porównywać jabłka i gruszki? O danych bibliometrycznych w różnych dziedzinach nauki”, *Nauka* 2/2013, 37-46, szczególnie str. 42-44. Autorzy pokazują, że AIS bierze pod uwagę przede wszystkim cytowania w czasopiśmie już wcześniej licznie cytowanych. Nie można więc szybko wyśrubować tego parametru dla danego czasopisma. Co więcej jest ono faktycznie normalizowane ze względu na poszczególne nauki i jest *public domain*, ale jest dostępne w ramach bazy czasopism Web of Science. Zmiana parametru oceny czasopism zepchnęło by w dół tabeli różne dziwne czasopisma o nikłym prestiżu a wysokim IF (por. cytowany artykuł). Nowsza analiza wraz z konkretnymi ciekawymi propozycjami zawarta jest w pracy J. Zakrzewski, K. Życzkowski, „Ministerstwo, punkty i artykuły naukowe (lub czasopisma)”, *Pauza Akademicka* 277-279 (2014/2015), str. 6.
2. Większa różnica punktowa między „najlepszymi” a najslabszymi czasopismami (tu zgadzam się z prof. Białynickim-Birulą). Ten postulat można by połączyć z zasadą, że grupa najwyżej ocenianych czasopism (wzorcowych) w każdej dziedzinie jest wybierana przez ekspertów.
3. W różnych dziedzinach nauki należałoby przyjąć odrębne regulacje w sprawie oceny publikacji wieloautorskich – w matematyce może zwykle dzielenie przez liczbę autorów byłoby w porządku, ale jednak tak aby publikacja współdzielona z 2 innymi w *Annals of Math.* była więcej warta niż dwie samodzielne publikacje w poślednim czasopiśmie. Trzeba mieć jednak świadomość, że w fizyce czy astronomii są publikacje z setkami autorów (co oznacza, że proste dzielenie w tamtych dziedzinach jest nie do przyjęcia).
4. Postulowałbym zwiększenie roli eksperckiej oceny 10 najważniejszych osiągnięć jednostki – to mógłby być substytut oceny eksperckiej całości osiągnięć. Faktycznie może warto by zmniejszyć liczbę do 5-6. Niech oceniana jednostka napisze czym są jej najważniejsze osiągnięcia naukowe. I niech eksperci to porównają. Tych światowych osiągnięć jest zwykle niewiele jeśli w ogóle.
5. Coś trzeba by zrobić z oceną niejednorodnych jednostek – nie wiem co?? Trzeba jednak rozwiązać problem, o którym pisze w punkcie 8 części pierwszej. To sprawa kluczowa jeśli chcemy uniknąć dziwnych wyników kolejnej oceny. Dla matematyki jest to szczególnie istotne, ze względu na inny wzorec cytowania i charakterystyk bibliometrycznych niż w naukach eksperymentalnych. W obecnym systemie zwykle każda „domieszka” np. fizyki winduje jednostkę matematyczną w górę.
6. Podoba mi się propozycja prof. Białynickiego-Biruli, aby szczegółowa punktacja czasopism oraz inne dokładne parametry były podane do wiadomości dopiero po zakończeniu procesu (dodajmy powinny być jednakowe dla wszystkich ocenianych publikacji w danej ocenie bez rozróżnienia na rok wydania – przeciwnie niż to było w ostatniej parametryzacji). Argumentem za takim podejściem jest uświadomienie sobie,

że mierzymy nie to co naprawdę chcemy (poziom naukowy) tylko mierzymy pośrednie dowody wysokiego poziomu badań. Wobec tego argument, że zasady powinny być znane z góry jest chybiony – zasada jest, że ma być wysoki poziom nauki, a jak to będzie mierzone jest już drugorzędne i należy do kompetencji oceniającej instytucji. Tu jednak chyba opuszczam grunt propozycji realistycznych...

Na koniec – sprawa jest poważna i środowisko matematyczne powinno zabrać głos – ze względu na terminarz jeśli ma być to głos Komitetu Matematyki, musiałoby to być na lutowym posiedzeniu KM. Jeśli uchwała ma mieć jakikolwiek realny wpływ, musi być konkretna i krótka oraz zawierać dobrze umotywowane propozycje konkretnych zmian. Nie bardzo widzę inne ciało, które reprezentuje całość matematyki i jest umocowane prawnie. Więc jak KM tego nie zrobi, to nikt tego nie zrobi. Lepiej gdybyśmy zdołali wypowiedzieć się w choć w kilku kluczowych sprawach niż zrezygnowali z tego na skutek braku zgody w wielu innych sprawach.

Paweł Domański