

DOROTA KAMIŃSKA^{A–D}, MARIUSZ KUSZTAŁ^{A, D–F}

Świadome i nieświadome zniekształcanie wyników badań naukowych

Intentional and Non-Intentional Distortion of Research Findings

Katedra i Klinika Nefrologii i Medycyny Transplantacyjnej, Uniwersytet Medyczny im. Piastów Śląskich we Wrocławiu, Wrocław, Polska

A – koncepcja i projekt badania; B – gromadzenie i/lub zestawianie danych; C – analiza i interpretacja danych;
D – napisanie artykułu; E – krytyczne zrecenzowanie artykułu; F – zatwierdzenie ostatecznej wersji artykułu

Streszczenie

W obecnych czasach mamy do czynienia z zalewem informacji naukowych i ryzykiem zniekształcenia wyników badań. Autorzy artykułu zwracają uwagę na częste pułapki wnioskowania statystycznego w badaniach biomedycznych. Świadome (zamierzone) i nieświadome zniekształcenie wyników bywa pochodną niewłaściwego planowania badań, zbyt małej grupy badanej, stosowania niewłaściwego sposobu analizy danych. Przesadna motywacja badaczy do publikowania wszelkich wyników badań z osiąganiem za wszelką cenę poziomu istotności statystycznej (np. $p < 0,05$) może z kolei skłaniać ich do nadmiernej kreatywności czy manipulacji danymi (np. uelastycznianie protokołu badania i rozszerzanie kryteriów włączenia, sprawdzanie/zmianie grup kontrolnych lub zmiana testów statystycznych) (**Polim. Med. 2015, 45, 1, 7–9**).

Słowa kluczowe: wnioskowanie statystyczne, istotność statystyczna, wyniki badań naukowych.

Abstract

The flood of scientific information increases the risk of research findings distortion. The authors brought attention to the possible pitfalls of statistical inference in biomedical research. Due to intentional or non-intentional errors in study design and size, data collection and inappropriate statistical analysis some of research findings are false or exaggerated. Researchers excessively motivated to publish all possible study results sometimes seek any possible way to achieve p-value less than 0.05. False-positive study results can be obtained in underpowered studies, excessive flexibility in study designs and participants recruitment, probing different control groups, creating special subgroups or changing statistical tests (**Polim. Med. 2015, 45, 1, 7–9**).

Key words: statistical inference, p value, statistical analysis.

W obecnych czasach mamy do czynienia z zalewem informacji naukowych. W morzu artykułów naukowych czasem trudno wyłonić ich obiektywną wartość. Niekiedy liczba publikowanych informacji dalece odbiega od ich jakości, tym bardziej że każdy z naukowców, zobowiązany do nieustannego wykazywania się osiągnięciami naukowymi, doświadcza na własnej skórze siły powiedzenia „publish or perish” („publikuj albo gin”).

Pozostaje więc pytanie, w jaki sposób nie dać się zaślepić kolejnymi doniesieniami wątpliwej jakości oraz – co ważniejsze – jak krytycznie podchodzić również do artykułów swojego autorstwa.

Prawie dekadę temu do rozpoczęcia dyskusji nad wiarygodnością, a nawet sensem prowadzenia badań naukowych i publikowania ich wyników, przyczynił się opublikowany w Plos Medicine esej Johna Ioannidisa o wdzięcznym tytule „Why Most Published Research Findings Are False” („Dlaczego większość opublikowanych badań naukowych jest nieprawdziwa”) [1]. Do tej pory przeciwnicy i zwolennicy tez przedstawionych w tej pracy dowodzą słuszności swych stanowisk.

Ioannidis po przeanalizowaniu wielu doniesień przedstawił następujące przyczyny nieświadomego zniekształcania wyników naukowych, czyli szanse na

zafałszowanie wyników: mała liczebność badanych prób, mniej wyraźny obserwowany efekt, większa liczba niedokładnie selekcjonowanych związków przyczynowo-skutkowych. Do tego dochodzi brak dokładności w definiowaniu efektów końcowych, zbyt duża elastyczność w sposobach analizy, wpływ zaangażowania finansowego lub światopoglądowego badacza oraz ogólna moda na dany temat badawczy.

Za Staniszem warto przypomnieć o pułapkach wnioskowania statystycznego, częstokroć nieuświadomionych lub wynikających z niewiedzy [2]. Podczas edukacji szkolnej i akademickiej jesteśmy uczeni sprawnego posługiwania się wnioskowaniem dedukcyjnym i indukcyjnym. Nadużywanie wnioskowania indukcyjnego prowadzi do tworzenia wniosków „ostatecznych”, niebudzących żadnych wątpliwości. Takie podejście jest sprzeczne z wnioskowaniem statystycznym, które *per se* zawiera element niepewności.

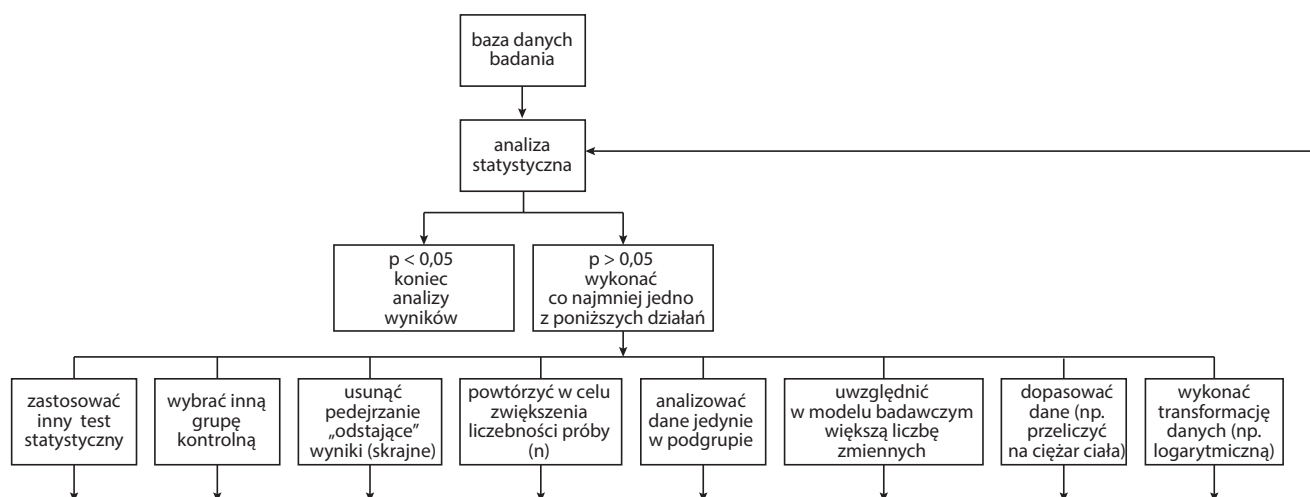
Znawcy tematu za jedną z najczęstszych pułapek przy wnioskowaniu statystycznym wymieniają zbyt małą liczebność grupy badanej – im jest ona mniejsza, tym większa jest niepewność wyciąganych wniosków. Wnioski wyciągane na podstawie zbyt małej próby mogą być całkowicie odmienne jakościowo od wniosków formułowanych na podstawie wystarczająco licznej próby. W rozwiązaniu tego problemu, który powinien być poruszony już na etapie planowania badania, pomocne okazują się programy statystyczne oferujące kalkulacje wielkości próby (np. Medcalc, SAS). Można też posłużyć się kalkulatorami specjalnie stworzonymi do tego celu, np. nQuery Advisor, PASS, „Power and Precision” (komercyjne) lub skorzystać z nieodpłatnych serwisów on-line, np. na stronach University of Iowa (<http://homepage.stat.uiowa.edu/~rlenth/Power/>) [3].

Kolejną, często spotykaną pułapką wnioskowania statystycznego jest ekstrapolacja wniosków poza badaną populację. Poprawna ekstrapolacja wniosków na szerszą populację w stosunku do populacji badanej wymaga niestety przeprowadzenia powtórnego wnioskowania na próbie pobranej z rozszerzonej populacji. Inna pułapka dotyczy utożsamiania korelacji między obserwacjami ze związkiem przyczynowo-skutkowym.

Dążenie do pierwszeństwa publikacji wyników na świecie (a przynajmniej we własnym kraju) prowadzi do braku powtarzania danych eksperymentów przez inne zespoły, a tym samym do niemożności ich obiektywnej weryfikacji. Dążenie badaczy do uzyskania poziomu istotności statystycznej (p zwykle $< 0,05$) za wszelką cenę prowadzi do zaniedbywania negatywnych wyników oraz do publikowania hipotez fałszywie dodatnich.

Niewłaściwe zastosowanie lub też próbowanie różnych testów statystycznych, dowolne usuwanie wartości skrajnych tylko w celu osiągnięcia pożądanej „istotności statystycznej” może prowadzić do świadomego zniekształcania wyników badań, a nawet ich zafałszowań. Innymi działaniami ukierunkowanymi na „zmniejszenie” wartości p są: sztuczne zwiększanie liczebności próby, analizy jedynie w podgrupach z małym rozrzutem pomiarowym, „dostrajanie” i ważenie danych przez dodawanie nowych współczynników (np. przeliczanie na ciężar lub powierzchnię ciała), transformowanie danych (np. do skali logarytmicznej), a nawet zmiany grupy kontrolnej (ryc. 1).

Według Ioannidisa nawet 86% opublikowanych wyników nie jest prawdziwa. Jego szacunki zostały ostatnio zweryfikowane przez Jagera i Leeka [4], którzy po praktycznym przeanalizowaniu artykułów publikowanych w czołowych wydawnictwach medycznych (*The*



Ryc. 1. Działania ukierunkowane na osiągnięcie poziomu istotności statystycznej ($p < 0,05$) powodujące zniekształcenie wyników badań naukowych (na podstawie [6])

Fig. 1. Activities aimed at obtaining the level of statistical significance ($p < 0,05$) that distort scientific research results (according to [6])

Lancet, *The Journal of the American Medical Association*, *The New England Journal of Medicine*, *The British Medical Journal* i *The American Journal of Epidemiology*) w latach 2000–2010 wykazali, że artykuły bazujące na istotności statystycznej były fałszywe „tylko” w 14%. Dało to pole do kolejnych dyskusji biostatystyków na ten temat [4–5]. Jednak problem bezmyślnej pogoni za $p < 0,05$ pozostaje aktualny. Publikowane są nawet rady, jak nie dać się zwieść pozornej istotności prezentowanych wyników [6]. West i Dupras proponują, aby

zwracać większą uwagę na istotność kliniczną danych zjawisk, analizę ryzyka, a nie tylko ryzyka względnego oraz ostrożną interpretację izolowanego istotnego p w przypadku zastosowania różnorodnych testów statystycznych. Podkreślają, że zjawiska nieistotne statystycznie mogą być istotne klinicznie [7]. Proponowana jest zmiana podejścia redaktorów czasopism, władz uczelni czy też komitetów naukowych przyznających fundusze na badania w kierunku większej promocji dużych, powtarzalnych badań naukowych [8].

Piśmiennictwo

- [1] Ioannidis J.P.A.: Why Most Published Research Findings Are False. *PLoS Medicine* 2005, 2(8) e124.
- [2] Stanisław A.: Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 1–3. Statystyki podstawowe. Kraków 2006.
- [3] <http://homepage.stat.uiowa.edu/~rlenth/Power/> dostęp dn.15.03.2015
- [4] Jager L.R., Leek J.T.: An estimate of the science-wise false discovery rate and application to the top medical literature. *Biostatistics* 2014, 15, 1, 1–12.
- [5] Gelman A., O'Rourke K.: Discussion: Difficulties in making inferences about scientific truth from distributions of published p-values. *Biostatistics* 2014, 15, 1, 18–23.
- [6] Motulsky H.J.: Common misconceptions about data analysis and statistics. *Pharma Res. Per.* 2014, 3(1), e00093.
- [7] West C.P., Dupras D.M.: 5 ways statistics can fool you — tips for practicing clinicians. *Vaccine* 2013, 31, 1550–1552.
- [8] Ioannidis J.P.A.: How to Make More Published Research True. *PLoS Medicine* 2014, 11(10) e1001747.

Adres do korespondencji:

Mariusz Kusztal
Klinika Nefrologii i Medycyny Transplantacyjnej
Uniwersytecki Szpital Kliniczny we Wrocławiu
ul. Borowska 213
50-556 Wrocław
Polska
e-mail: mariusz.kusztal@umed.wroc.pl

Konflikt interesów: nie występuje

Praca wpłynęła do Redakcji: 20.03.2015 r.

Zaakceptowano do druku: 20.03.2015 r.

Received: 20.03.2015

Accepted: 20.03.2015