



Poszukiwanie sensownej bazy dowodów empirycznych w psychiatrii

The quest for a meaningful evidence base in psychiatry

ROBIN EMSLEY, SUE HAWKRIDGE

Department of Psychiatry, Faculty of Health Science, Stellenbosch University, Tygerberg 7505, Cape Town, Afryka Południowa

Podczas gdy reszta środowiska lekarskiego szybko i śmiało wkroczyła w erę medycyny opartej na dowodach naukowych, w psychiatrii początkowo panowała niechęć do podążania w ślady kolegów i długo trwało, zanim przekonano się do zasad tego podejścia. Obecnie jednak coraz więcej psychiatrów entuzjastycznie odnosi się do podejścia opartego na dowodach naukowych i domaga się „dowodów” dla wszelkich interwencji terapeutycznych. Kłopot przy tym podejściu polega na tym, że dane naukowe często nie wskazują na jednoznaczny wynik, bywają niespójne, a nawet wzajemnie sprzeczne, co stwarza niebezpieczną sytuację – ci, którzy mają w tym interes, mogą sobie dobierać dowody naukowe według potrzeb.

Instytucje finansujące leczenie skwapliwie wyłapują badania, które nie wykazały wyższości nowszych farmaceutyków i na podstawie ich wyników przekonują do powrotu do (tańszych) leków przeciwpsychotycznych pierwszej generacji (PPG). Równocześnie zwolennicy nowszych leków pomijają te badania, wskazując na ich niedociągnięcia metodologiczne, z natury obecne we wszystkich klinicznych badaniach lekowych.

Ponieważ symptomatologia psychiatryczna jest w dużej mierze subiektywna, gromadzenie dowodów naukowych przy użyciu obiektywnych metod pomiaru jest tym trudniejsze. Wprawdzie nastąpił wiele obiecujący postęp i wzrosła nasza zdolność skutecznego leczenia większości zaburzeń psychicznych, lecz w praktyce klinicznej nadal istnieją duże braki, a wyniki leczenia „w świecie rzeczywistym” często są niezadowolające. Problem polega częściowo na zdefiniowaniu najlepszego sposobu postępowania w praktyce na podstawie dostępnych dowodów naukowych. Ilość „danych naukowych” publikowanych co miesiąc w specjalistycznych czasopismach bywa oszałamiająca.

Artykuł Fleischackera i Goodwina w samą porę przedstawia wnikliwe omówienie niektórych trudności, jakie napotykają psychiatrzy próbując przełożyć wyniki badań naukowych na najlepszą praktykę. Tak więc, randomizowane badania kontrolowane [RCTs] – kamień węgielny medycyny opartej na dowodach naukowych – są obiektem ataków w psychiatrii, krytykuje się je między innymi za to, że niedokładnie odzwierciedlają warunki „świata rzeczywistego” [1]. Wysokie i rosnące wskaźniki odpowiedzi na placebo oraz przerywania leczenia w randomizowanych badaniach kontrolowanych stały się nawracającym koszmarem statystyków, budząc poważne wątpliwości co do trafności wyników tych badań. Próbuując przeciwdziałać owym wadom RCTs podejmuje się tak zwane badania „pragmatyczne”, coraz częściej publikowane w literaturze. Chociaż noszą one kuszące i wiele obiecujące nazwy skrótowe, takie, jak CATIE [2], CUtLASS [3], CAFE [4], EUFEST [5] i STAR*D [6], jednak istnieje obawa, że badania te jeszcze

bardziej zamącą obraz. Okazują się równie trudne do interpretacji i wywołują wiele kontrowersji. Wydaje się, że przeprowadzenie każdego nowego badania generuje kolejną grupę pytań, na które nie ma odpowiedzi.

Fleischacker i Goodwin opowiadają się za utrzymaniem zarówno RCTs, jak i badań pragmatycznych, przy czym te ostatnie powinno się przeprowadzać na wcześniejszym etapie opracowywania leku, zanim opinie na jego temat utrwalą się. Dzięki połączeniu zalet naukowo rygorystycznych RCTs i pragmatycznych badań bliższych praktyce w świecie rzeczywistym możemy mieć nadzieję, że przybliżymy się do ustalenia, jakie sposoby leczenia są najlepsze dla naszych pacjentów. Chociaż niewątpliwie ma to sens, samo w sobie może nie wystarczyć, ponieważ trudność polega nie tylko na uzyskaniu dowodów (danych empirycznych), ale także na ich interpretacji. Problem częściowo może polegać na tym, że mamy za duże oczekiwania wobec każdej poszczególnych próby lekowej – chociaż badania te zwykle zaprojektowane są tak, by dawały odpowiedź na jedno czy dwa pytania, mimo to często próbujemy ekstrapolować ich wyniki na inne zagadnienia i inne populacje pacjentów. Na przykład, w programie CATIE stwierdzono, że perfenazyna, lek przeciwpsychotyczny pierwszej generacji (PPG), wykazuje zdumiewająco dobre działanie w porównaniu z lekami przeciwpsychotycznymi drugiej generacji. Uogólnienie tego wniosku na inne PPG, takie, jak haloperidol, jest jednak potencjalnie niebezpieczne. Żadne pojedyncze badanie kliniczne, czy to randomizowane kontrolowane, czy pragmatyczne, nie jest zaprojektowane tak, by mogło dać odpowiedź na wszystkie pytania lub stanowiło podstawę do sformułowania definitywnego standardu leczenia – każde badanie dodaje coś nowego do podstaw naszej wiedzy (istniejącej bazy danych). Oznacza to, że „dowody naukowe”, na których opiera się praktyka, to duże zbiory niekiedy niespójnych wiadomości. Klinicyści muszą być w stanie wprowadzić dające się uzasadnić wnioski z każdego nowego elementu wiedzy w swęj codziennej praktyce, zaś trafna interpretacja i przekład będą wymagały podejścia kumulatywnego na długą metę, nie zaś pochopnego i niekiedy oportunistycznego koncentrowania się na każdym nowym fragmencie pojawiających się danych.

Sądźmy jednak, że ostrożność w podchodzeniu do *status quo* nie wystarcza; musimy zrobić coś więcej, aby wyeliminować tendencyjność. Kwestią, którą trzeba się pilnie zająć, jest tendencyjność publikacji. Konsekwencje wybiórczego publikowania pozytywnych wyników badań stały się boleśnie jasne w kontrowersjach dotyczących stosowania leków przeciwdepresyjnych i innych pokrewnych farmaceutyków u młodzieży [7]. W różnych niedawno ustanowionych rejestrach badań klinicznych należałoby dołożyć starań, aby nie dopuścić do powtórzenia się takiej mianowicie sytuacji, że przyjęte i wdrożone międzynarodowe standardy leczenia

oparto, jak się wydaje nieświadomie, na niepełnych danych. Wpływ firm farmaceutycznych i produkujących sprzęt medyczny, które oskarża się o narażanie na szwank uczciwości zawodowej środowiska medycznego, wymaga starannej kontroli [8, 9].

Jakość danych klinicznych pochodzących z RCTs można poprawić poprzez minimalizowanie zachęt do udziału w badaniach [zwłaszcza przy szybkiej rekrutacji do badań], zapewnienie nieskazitelnej metodologii, oraz stosowanie odpowiednich miar wyników leczenia. Rygorystyczne szkolenie badaczy może zwiększyć precyzję zbieranych danych klinicznych. Użyteczność badań pragmatycznych można zwiększyć uwzględniając szersze spektrum kontekstów klinicznych, włączając populacje pacjentów z krajów rozwijających się oraz innych środowisk o niskich dochodach. Co być może najważniejsze, klinicyści będą musieli reprezentować nie-dogmatyczne podejście, mieć gruntowną wiedzę o *wszystkich* dowodach naukowych oraz zdrowy rozsądek kliniczny, nadal niezastąpiony.

PIŚMIENNICTWO

1. Gilbody S, Wahlbeck K, Adams C. Randomized controlled trials in schizophrenia: a critical perspective on the literature. *Acta Psychiatr Scand* 2002; 105: 243–51.
2. Lieberman JA, Stroup TS, McEvoy JP et al. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl J Med* 2005; 353: 1209–23.
3. Jones PB, Barnes TR, Davies L et al. Randomized controlled trial of the effect on Quality of Life of second- vs first-generation antipsychotic drugs in schizophrenia: Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1). *Arch Gen Psychiatry* 2006; 63: 1079–87.
4. McEvoy JP, Lieberman JA, Perkins DO et al. Efficacy and tolerability of olanzapine, quetiapine, and risperidone in the treatment of early psychosis: a randomized, double-blind 52-week comparison. *Am J Psychiatry* 2007; 164: 1050–60.
5. Kahn RS, Fleischhacker WW, Boter H et al. Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomised clinical trial. *Lancet* 2008; 371: 1085–97.
6. Trivedi MH, Rush AJ, Wisniewski SR et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry* 2006; 163: 28–40.
7. Hammad TA, Laughren T, Racoosin J. Suicidality in pediatric patients treated with antidepressant drugs. *Arch Gen Psychiatry* 2006; 63: 332–9.
8. DeAngelis CD, Fontanarosa PB. Impugning the integrity of medical science: the adverse effects of industry influence. *JAMA* 2008; 299: 1833–5.
9. Fava GA. Financial conflicts of interest in psychiatry. *World Psychiatry* 2007; 6: 19–24.



Postępy Psychiatrii i Neurologii 2009; 18 (2): 124–125

Rola badań skuteczności i efektywności

A. John Rush

Duke-NUS Graduate Medical School Singapore, Jalan Bukit Merah, 169547 Singapur

Fleischhacker i Goodwin zwracają uwagę na kilka istotnych kwestii dotyczących roli badań nad skutecznością i efektywnością, w kontekście przeglądu najnowszych dużych badań pragmatycznych. Na początku sugerują, że gdyby badania skuteczności i efektywności doprowadziły do rozbieżnych wniosków, powinniśmy ufać badaniom skuteczności. Odnotowują również, że wyniki badań skuteczności, jeśli są pozytywne, mogą odnosić się do ograniczonej podgrupy pacjentów (tzn. do osób spełniających często wyśrubowane kryteria kwalifikacji stosowane w badaniach skuteczności, lub leczonych „w warunkach badawczych”, czyli otrzymujących duże dawki leku). Później we wnioskach są jednak bardziej łaskawi dla badań efektywności, powiadając, że są one potrzebne i proponując nawet, by domagać się badań efektywności zanim producent otrzyma zezwolenie i dany specyfik będzie w pełni dostępny na rynku. Więc jak to jest? Czy badania efektywności są przydatne, czy nie?

W pewnym stopniu zgadzam się z nimi, ale nie do końca. Uważam, że nie ma dwojakiego rodzaju badań (tzn. skuteczności w odróżnieniu od efektywności). Projektuje się określone plany badawcze, aby uzyskać odpowiedź na określone pytania. Z różnych pytań wynikają różne plany

badawcze. Z różnych planów badawczych uzyskamy różne odpowiedzi, tak, jak być powinno, ponieważ mają odpowiedzieć na różne pytania.

Badania skuteczności (faza II–III) projektuje się z maksymalną trafnością wewnętrzną, aby odpowiedzieć na pytania w rodzaju: jaka jest skuteczność, bezpieczeństwo i tolerancja leczenia „X” w porównaniu z placebo, tzn. aby wyodrębnić kliniczny wpływ *samej tylko* cząsteczki chemicznej lub urządzenia na pacjenta (czego wyrazem są skutki uboczne) i na chorobę (czego wyrazem jest efekt terapeutyczny lub pogorszenie).

Badania efektywności obejmują mnóstwo różnych planów badawczych, które mają przynieść odpowiedź wiele różnych pytań. Ich specyfika, jak zauważyli Fleischhacker i Goodwin, polega na tym, że badana jest szersza populacja pacjentów, używa się większej liczby „klinicznie istotnych” miar wyników leczenia, i prowadzi się leczenie w warunkach „zwykłych” w odróżnieniu od „badawczych” (co może, ale nie musi zwiększać liczbę osób utrzymujących się w badaniu i wiąże się z ryzykiem zaniżania dawek leku). Kiedy jakieś badanie skuteczności wykaże skuteczność leku, całkiem prawdopodobne, że wielkość tego efektu w praktyce będzie inna, zależnie od tego, *kto* jest leczony