



Medycyna oparta na dowodach naukowych: blaski i cienie

Jacek Jassem

Gdański Uniwersytet Medyczny

VIII Letnia Akademia Onkologiczna dla Dziennikarzy

*W sytuacji braku aktualnych
i wiarygodnych publikacji,
lekarz ma wiele możliwości.*

*To właśnie sprawia,
że medycyna jest w równej mierze
nauką i sztuką.*

D. Isaacs i D. Fitzgerald

Evidence-based medicine

- praktyka medyczna oparta
na wynikach wiarygodnych
i aktualnych badań naukowych

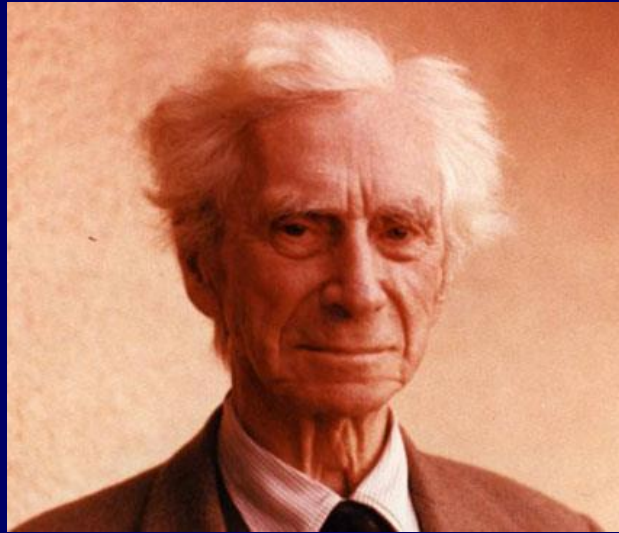
Evidence-based medicine (EBM)

- Integracja najlepszych dowodów naukowych z doświadczeniem klinicznym i oczekiwaniami pacjentów
- Złoty standard praktykowania i nauczania medycyny klinicznej

Cele *evidence-based medicine*

Oddzielenie:

- dowodów od propagandy
- pewności od przypuszczeń
- wyników badań od zapewnień
- racjonalizmu od przesądów
- nauki od folkloru
- wiedzy od dogmatów



*Powszechne wprowadzenie zwyczaju
potwierdzania przekonań dowodami
i przypisywanie tym ostatnim odpowiedniego
stopnia pewności, pozwoliłoby pokonać
większość chorób trapiących świat*

Bertrand Russell, 1957

Tradycyjne rozwiązywanie problemu klinicznego

Własne obserwacje kliniczne



Analiza zachodzących procesów fizjologicznych
i patologicznych



Zapoznanie się z wiedzą podręcznikową



Zasięgnięcie opinii eksperta

Wady związane z tradycyjnym (opartym na „zdrowym rozsądku”) rozwiązywaniem problemów klinicznych

- Powszechne uznanie i stosowanie metod nieskutecznych lub nawet szkodliwych
- Rozbieżne opinie ekspertów (wówczas istnieje „najlepsza” metoda postępowania, a w przypadku niezgodnych stanowisk, jedna ze stron ma zawsze rację)
- Opóźnienie w rozpowszechnieniu wiarygodnej wiedzy klinicznej
- Brak zrozumienia koncepcji i pojęć dla przekazywania tej wiedzy

Nowoczesne rozwiązywanie problemu klinicznego

Sformułowanie pytania klinicznego



Znalezienie i wybór wiarygodnych badań na ten temat



Zapoznanie się z ich wynikami



Ocena ich wiarygodności



Integracja nowych wiadomości z posiadaną wiedzą



Podjęcie decyzji

Czy wiedza „podręcznikowa” nie jest wystarczająca?

Nabywana przez nas wiedza w dziedzinie medycyny
w ciągu 10 lat dezaktualizuje się w 50%.

W przypadku, kiedy nie jest ona
na bieżąco uzupełniana, w sposób nieuchronny
jej średni poziom stale się obniża

Jakość dowodów klinicznych

Poziom

Źródło dowodów

- I Dowody pochodzące z prawidłowo zaplanowanych i przeprowadzonych badań klinicznych z losowym doбором chorych lub metaanalizy badań klinicznych z randomizacją
 - II Dowody pochodzące z prawidłowo zaplanowanych i przeprowadzonych prospektywnych badań obserwacyjnych (badania kohortowe bez losowego doboru)
 - III Dowody pochodzące z retrospektywnych badań obserwacyjnych lub kliniczno--kontrolnych
 - IV Dowody pochodzące z doświadczeń uzyskanych w klinicznej praktyce lub/i opiniach ekspertów
-

Badanie kliniczne III fazy: podstawa *evidence based medicine*

- badanie porównawcze
z losowym doborem chorych;
najbardziej obiektywna metoda
oceny nowych metod
diagnostycznych i leczniczych

Ograniczenia zastosowania badań z randomizacją w praktyce klinicznej

- Metodologia badań z randomizacją bywa bezużyteczna (np. „nieżyciowe” kryteria doboru)
- Wyniki badań z randomizacją stanowią „przeciętną” dla całej badanej grupy i w odniesieniu do pojedynczego chorego mogą być źródłem błędnych decyzji
- Niektóre problemy nie zostały jeszcze zweryfikowane badaniami z randomizacją (ale brak dowodów skuteczności nie stanowi dowodu nieskuteczności)

**Warunkiem upowszechnienia
wyników badań jest
ich opublikowanie.
Czy jednak tak się dzieje?**

Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy

Erick H. Turner, M.D., Annette M. Matthews, M.D., Eftihia Linardatos, B.S.,
Robert A. Tell, L.C.S.W., and Robert Rosenthal, Ph.D.

- Spośród 74 badań w rejestrze FDA, 31% badań (z udziałem 3449 chorych) nigdy nie opublikowano
- Opublikowano 36 spośród 37 badań „pozytywnych”
- Spośród 33 badań „negatywnych” lub „wątpliwych” 22 nie opublikowano, a 11 opublikowano jako „pozytywne”
- Przegląd opublikowanych badań sugeruje, że 94% było „pozytywnych” (w rzeczywistości 54%)

Factors Associated With Failure to Publish Large Randomized Trials Presented at an Oncology Meeting

- 510 streszczeń zjazdowych dotyczących badań III fazy przedstawionych w latach 1989-1998 na kongresie ASCO
- 26% spośród nich nie zostało opublikowanych po 5 latach, w tym 19% „pozytywnych” i 32% „negatywnych” ($p < 0,001$)

Commentary: Practicing on the Tip of an Information Iceberg?
Evidence of Underpublication of Registered Clinical
Trials in Oncology

SCOTT RAMSEY,^{a,b} JOHN SCOGGINS^c

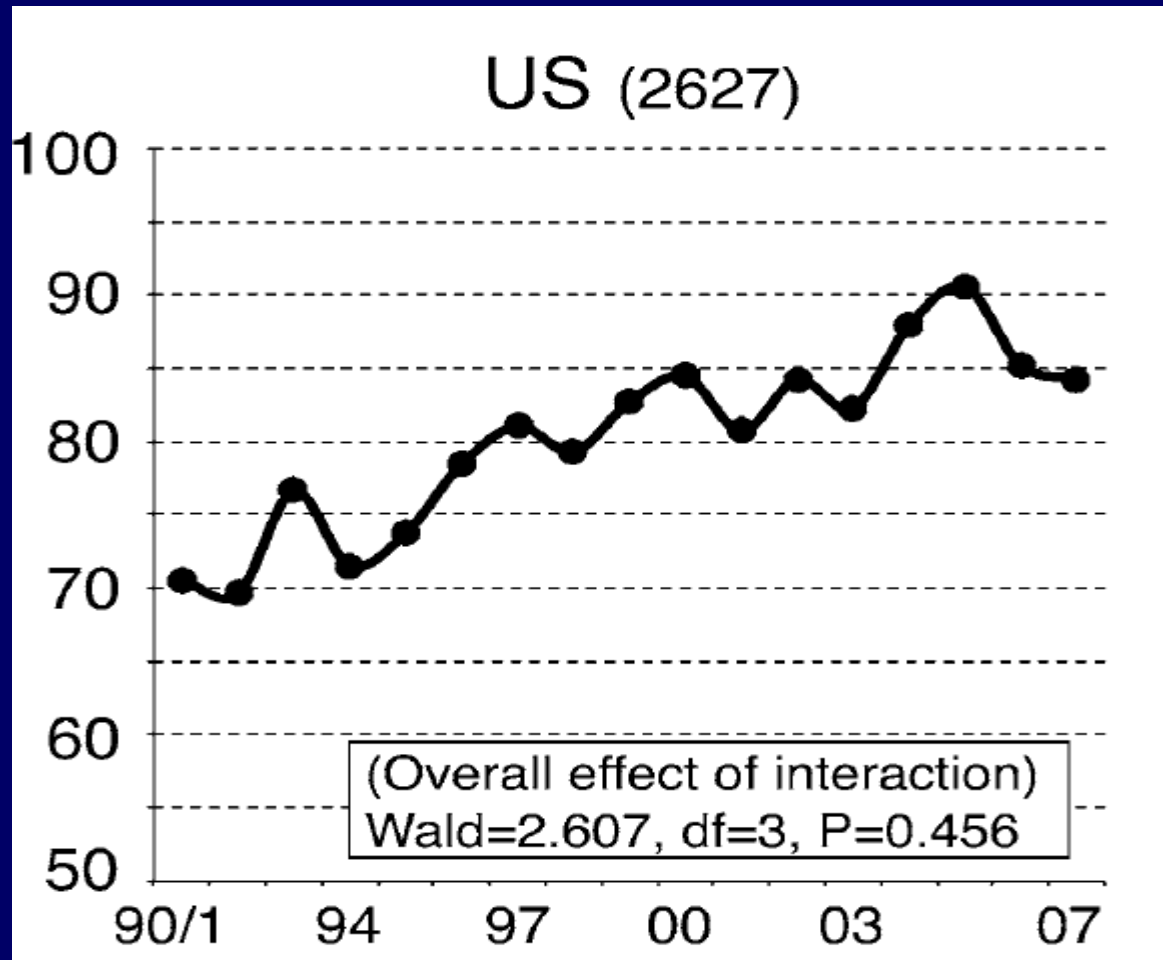
- Spośród 2.028 zarejestrowanych przez NIH badań klinicznych w onkologii, opublikowano 17,6%, w tym 59% badań akademickich i 5,9% sponsorowanych przez firmy farmaceutyczne
- Spośród opublikowanych badań 65% było „pozytywnych”, w tym 50% akademickich i 79% sponsorowanych

Publication bias in clinical trials due to statistical significance or direction of trial results (Review)

Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K

- Systematyczny przegląd badań III fazy z baz MEDLINE, EMBASE, Ovid i Science Citation Index
- Wyniki klasyfikowane jako pozytywne ($P < 0.05$) lub negatywne
- Szanse opublikowania badania pozytywnego było czterokrotnie większe niż badania negatywnego
- Średni czas do opublikowania badań pozytywnych wynosił 4-5 lat, a negatywnych 6-8 lat

Udział „pozytywnych” badań wśród ogółu publikacji



Prawdopodobne przyczyny *publication bias*

- Brak zapału do publikowania „negatywnych” badań przez ich autorów i sponsorów
- Wzrastające współzawodnictwo o granty, zachęcające do publikowania „pozytywnych” badań
- Mniejsze szanse przyjęcia do druku doniesień zawierających „negatywne” wyniki (zarówno z powodu niechętnego podejścia recenzentów, jak i redaktorów czasopism)

Zagrożenia związane z *publication bias*

Aspekt etyczny:

- chorzy poddają się badaniu wierząc, że przyczyni się to do skuteczniejszego leczenia przyszłych chorych
- środowisko naukowe oczekuje upowszechnienia wyników

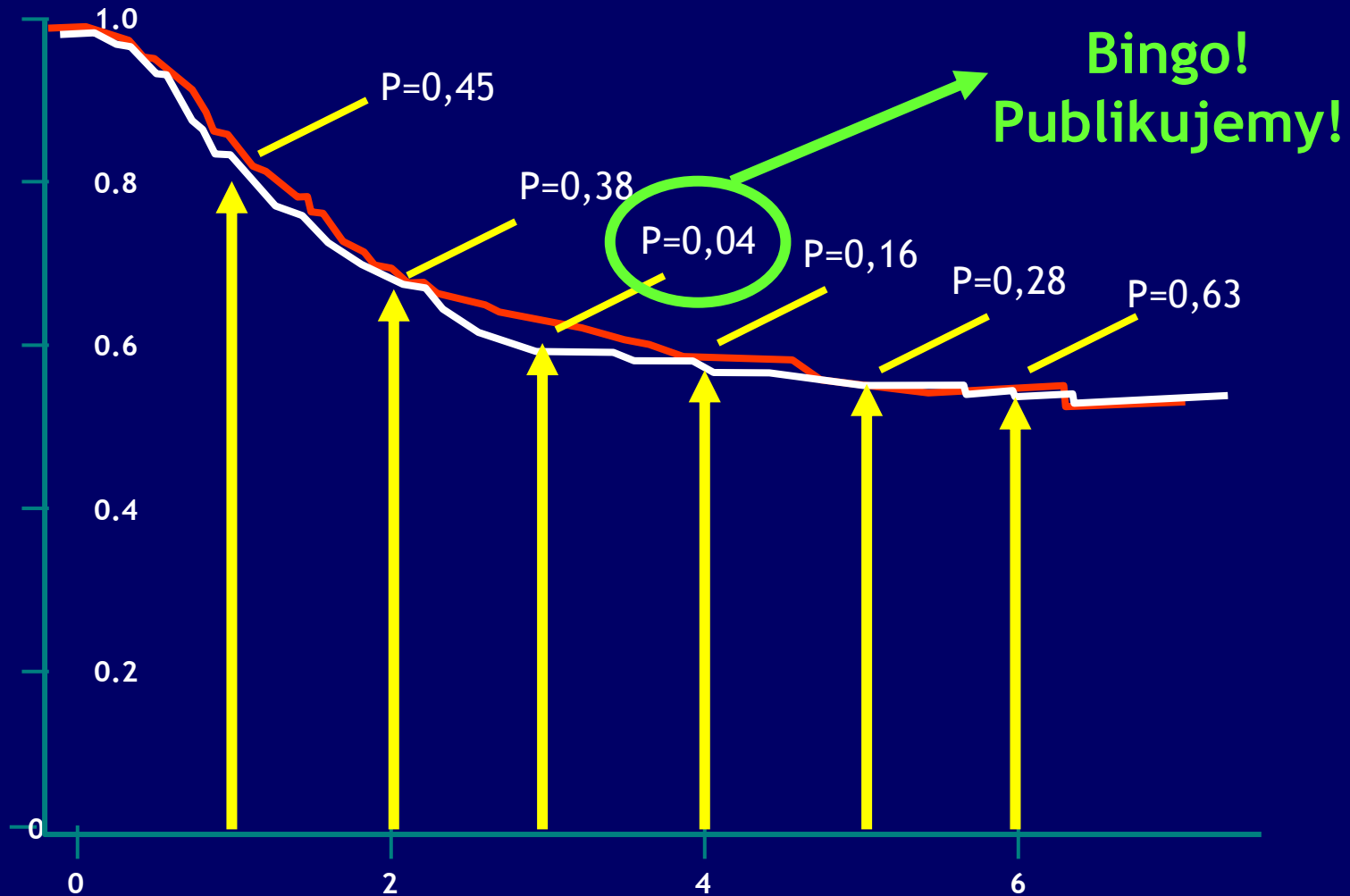
Aspekt naukowy:

- nieopublikowanie badania nie miałoby istotnego znaczenia, gdyby zjawisko to dotyczyło w tym samym stopniu badań „pozytywnych” i „negatywnych”; tak jednak nie jest

Publication bias in situ

- Wielokrotna analiza tych samych danych
- Wybór *ex post* ocenianych wskaźników
- Analiza podgrup
- „Masowanie” danych

Wielokrotne analizowanie wyników badania



Badanie ATAC: porównanie skuteczności dwu leków hormonalnych w raku piersi

ARTICLES

VOLUME 28 • NUMBER 7 • MARCH 1, 2008

Retrospective Analysis of Time to Recurrence in the ATAC Trial According to Hormone Receptor Status:

A

M

Pu

From the

A — The ATAC adjuvant breast cancer trial in postmenopausal women: baseline endometrial subprotocol data

a: **ONCOLOGY**

A lower incidence of gynecologic adverse events and interventions with anastrozole than with tamoxifen in the ATAC trial

Sean R. Duffy, MD; Wolfgang Distler, MD; Anthony Howell, MD; Jack Cuzick, PhD; Michael Baum, MD

OBJECTIVE: Results of the Arimidex, Tamoxifen, Alone or in Combination (ATAC) trial have shown that tamoxifen is associated with a significantly higher incidence of gynecologic adverse events than anastrozole.

including an almost 4-fold increase in the number of hysterectomies (5.1% vs 1.3%; $P < .0001$). The majority of the gynecologic adverse events with tamoxifen occurred during the first 2.5 years.

STUDY DESIGN: This was a retrospective analysis of all gynecologic

CONCLUSION: The lower incidence of gynecologic adverse events and

W żadnej z tych publikacji
nie pokazano krzywych
całkowitego przeżycia

...bo okazało się, że „stare”
i „nowe” leczenie
mają identyczną skuteczność

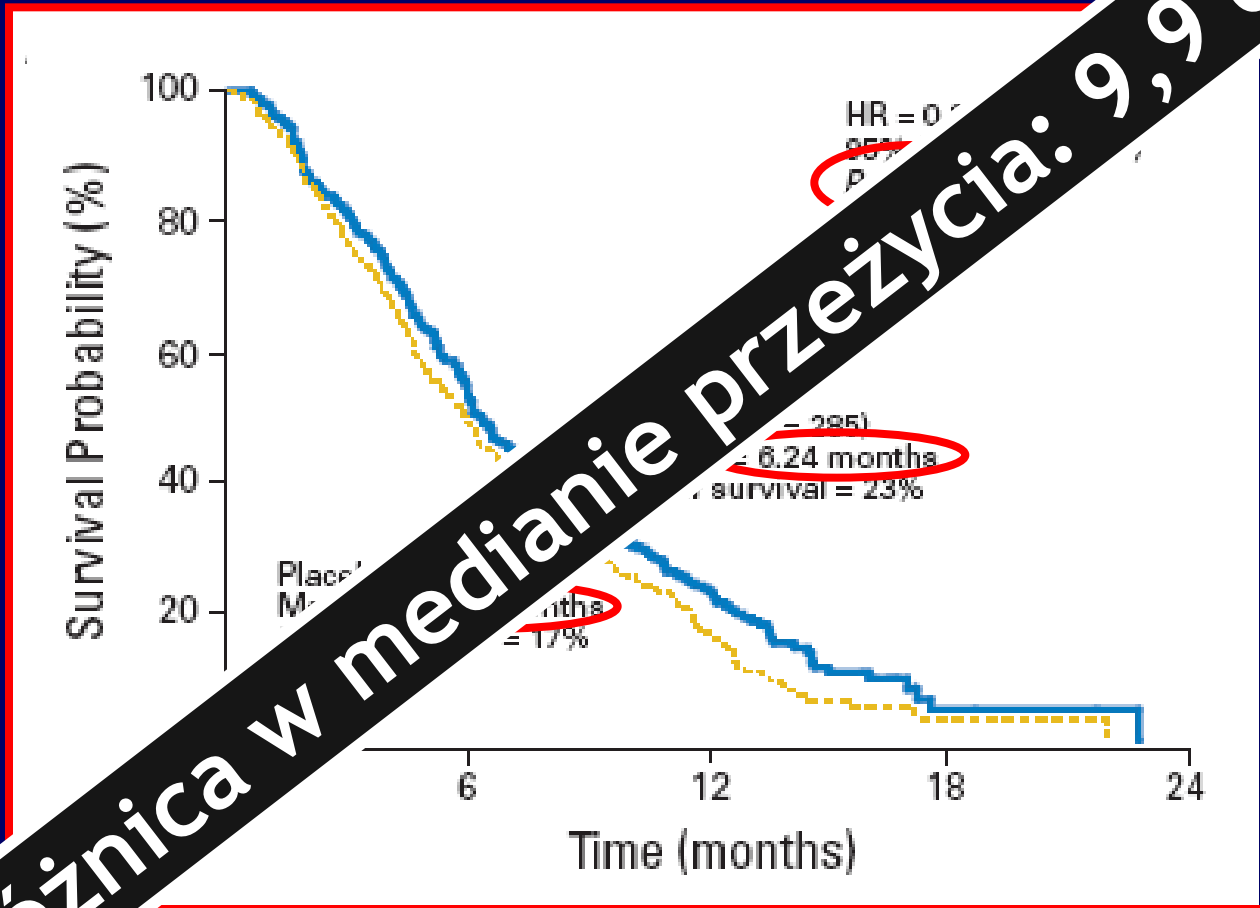
Up-Front Use of Aromatase Inhibitors As Adjuvant Therapy for Breast Cancer: The Emperor Has No Clothes

Bostjan Seruga and Ian F. Tannock, *Division of Medical Oncology and Hematology, Princess Margaret Hospital, Toronto, Ontario, Canada*



*"...im bardziej Kubuś zaglądał do studni,
tym bardziej Prosiaczka tam nie było" (Alan Milne)*

Czy każda „znamienna” różnica ma znaczenie kliniczne?

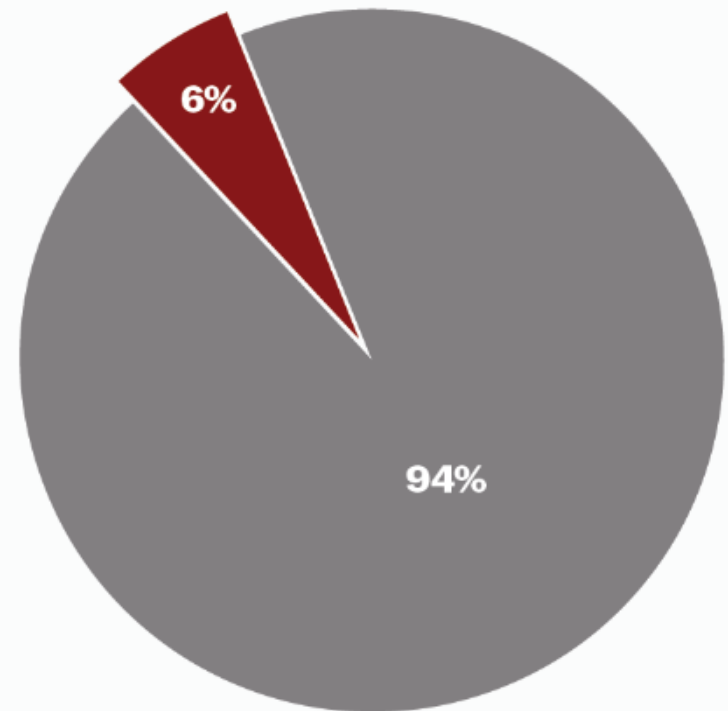


Różnica w medianie przeżycia: 9,9 dni

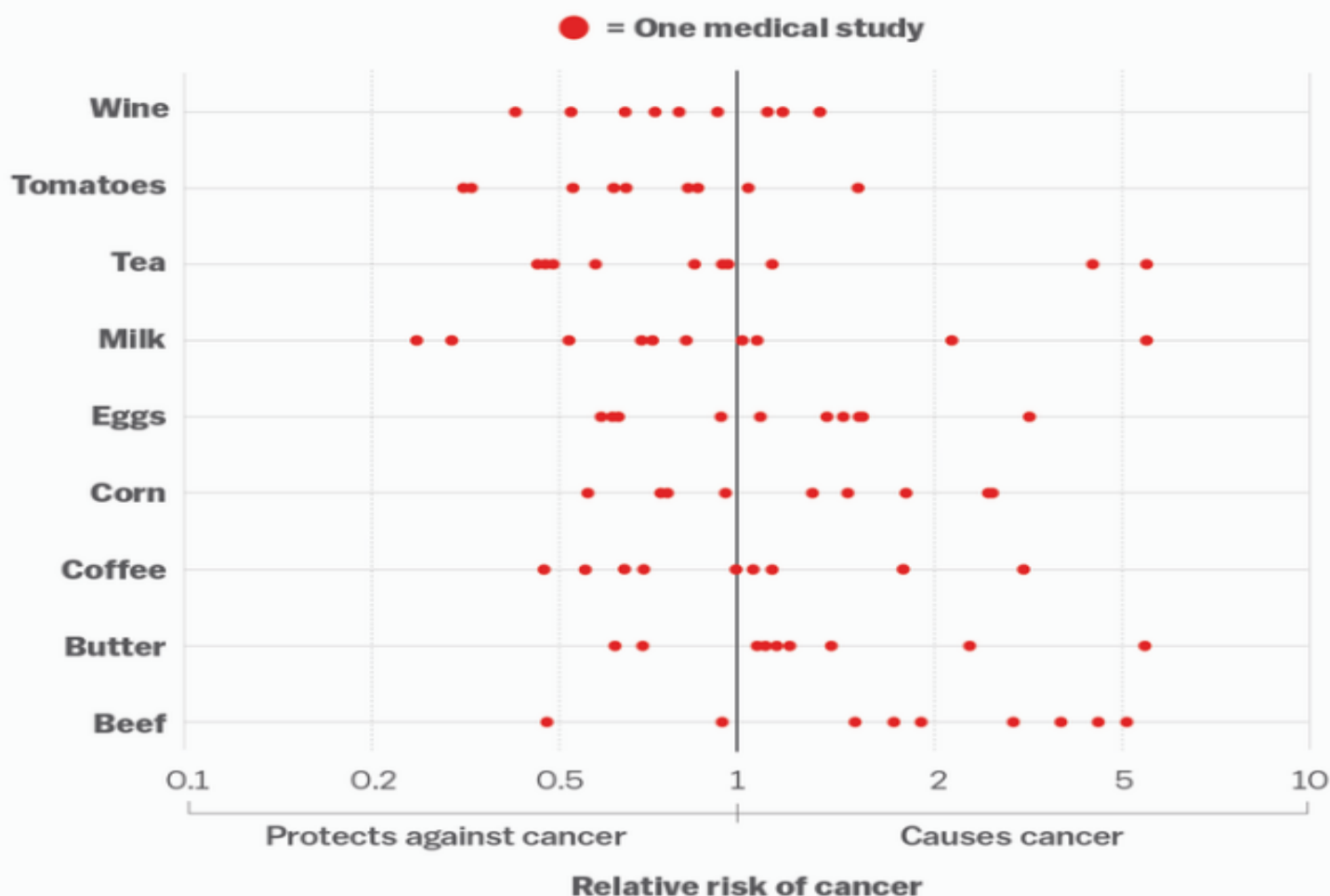
Dlaczego nie powinniśmy (prawie nigdy) wierzyć wynikom pojedynczych badań?

Spośród 50.000 nowych doniesień recenzowanych rocznie, tylko 3.000 (6%) jest na tyle wartościowych, aby zmienić praktykę kliniczną

SOURCE: Haynes, *Evidence Based Nursing*



Wszystko co jemy zarówno powoduje raka, jak i zapobiega rakowi



SOURCE: Schoenfeld and Ioannidis, *American Journal of Clinical Nutrition*

Jak można rozwiązać te problemy?

- Wprowadzić generalną zasadę publikowania *wszystkich* badań i *wszystkich* wyników, także tych negatywnych
- Zwiększyć wysiłki w poszukiwaniu niepublikowanych wyników w celu włączenia do metaanaliz wszystkich istotnych badań
- Wprowadzić bezwzględny wymóg przestrzegania przez wszystkie czasopisma medyczne zaleceń *International Committee of Medical Journal Editors* z 2004 roku, nakazujących rejestrację wszystkich badań klinicznych przed ich rozpoczęciem

Siedem alternatyw dla *evidence based medicine*

BMJ, 1998

Alternatywa 1: praktyka medyczna oparta na zasługach i własnym doświadczeniu (*eminence-based medicine*)

Im lekarz bardziej zasłużony i wyżej stojący w hierarchii, tym mniejszą wagę przypisuje czemuś tak banalnemu, jak dowody naukowe. Kolega taki pokłada wzruszającą wiarę w doświadczenie kliniczne, definiowane jako „powtarzanie tych samych błędów przez wiele lat z coraz większym przekonaniem”. Równoważy ono każdą liczbę wiarygodnych badań. Siwe włosy lub łysinę u owego lekarza zwie się „efektem aureoli”.

Alternatywa 2: praktyka medyczna oparta na głośności i porywczoci wypowiedzi (*vehemence-based medicine*)

Zastąpienie dowodów głośnym i porywczym wyrażaniem swych opinii jest skuteczną techniką uciszania co lękliwszych kolegów oraz przekonywania rodziny chorego o swych lekarskich umiejętnościach.

Alternatywa 3: praktyka medyczna oparta na elokwencji (*eloquence-based medicine*)

Wymowa powinna być równie gładka,
jak pozostałe atrybuty - opalone przez cały rok
ciało, goździk w butonierce, krawat
z jedwabiu i garnitur od Armaniego.

Elegancja w ubiorze i elokwencja są niezwykle
skutecznymi substytutami dowodów naukowych.

Alternatywa 4: praktyka medyczna oparta na zrządzeniu Opatrzności (*providence-based medicine*)

Kiedy troskliwy doktor nie ma pojęcia,
co zrobić, decyzję chętnie składa w ręce
Wszchemogącego. Niestety, zbyt wielu
lekarzy nie potrafi się oprzeć
tej pokusie.

Alternatywa 5: praktyka medyczna oparta na niepewności (*diffidence-based medicine*)

Jeden lekarz dostrzega problem, inny zaś nie tylko dostrzega, ale poszukuje rozwiązania. Lekarz ogarnięty niepewnością nie jest w stanie podjąć żadnej decyzji. Oczywiście niejednokrotnie bywa to lepsze niż zrobienie czegokolwiek wyłącznie po to, by nie zranić bezczynnością własnej lekarskiej dumy.

Alternatywa 6: praktyka medyczna oparta na strachu (*nervousness-based medicine*)

Strach przed pozwaniem do sądu silnie pobudza do wykonywania zbędnych badań diagnostycznych i stosowania niepotrzebnego leczenia. W atmosferze lęku jedynym uzasadnionym badaniem jest to, o którym się zapomniało.

Alternatywa 7: praktyka medyczna oparta na niezachwianej pewności siebie (*confidence-based medicine*)

Dotyczy jedynie chirurgów

Fenotypowe różnice pomiędzy internistami, chirurgami i gwiazdami filmowymi (badanie porównawcze)

Trilla, A. et al. BMJ 2006;333:1291-1293

Cele badania

Sprawdzenie hipotezy, że chirurdzy są średnio wyżsi i przystojniejsi niż interniści, oraz porównanie obu tych grup z aktorami, którzy grają role lekarzy

Miejsce badania

Typowy szpital uniwersytecki
w Hiszpanii, zlokalizowany
w Barcelonie, a nie na zapyztałej
prowincji

Uczestnicy

Losowo wybrana próbka 12 chirurgów i 12 internistów oraz 4 osoby stanowiące grupę kontrolną (aktorzy grający role lekarzy), dobrani wiekiem (około 50 lat) i płcią (wyłącznie mężczyźni)

Metody

Niezależny komitet składający się wyłącznie z kobiet, oceniający „wskaźnik dobrego wyglądu” w skali 1-7

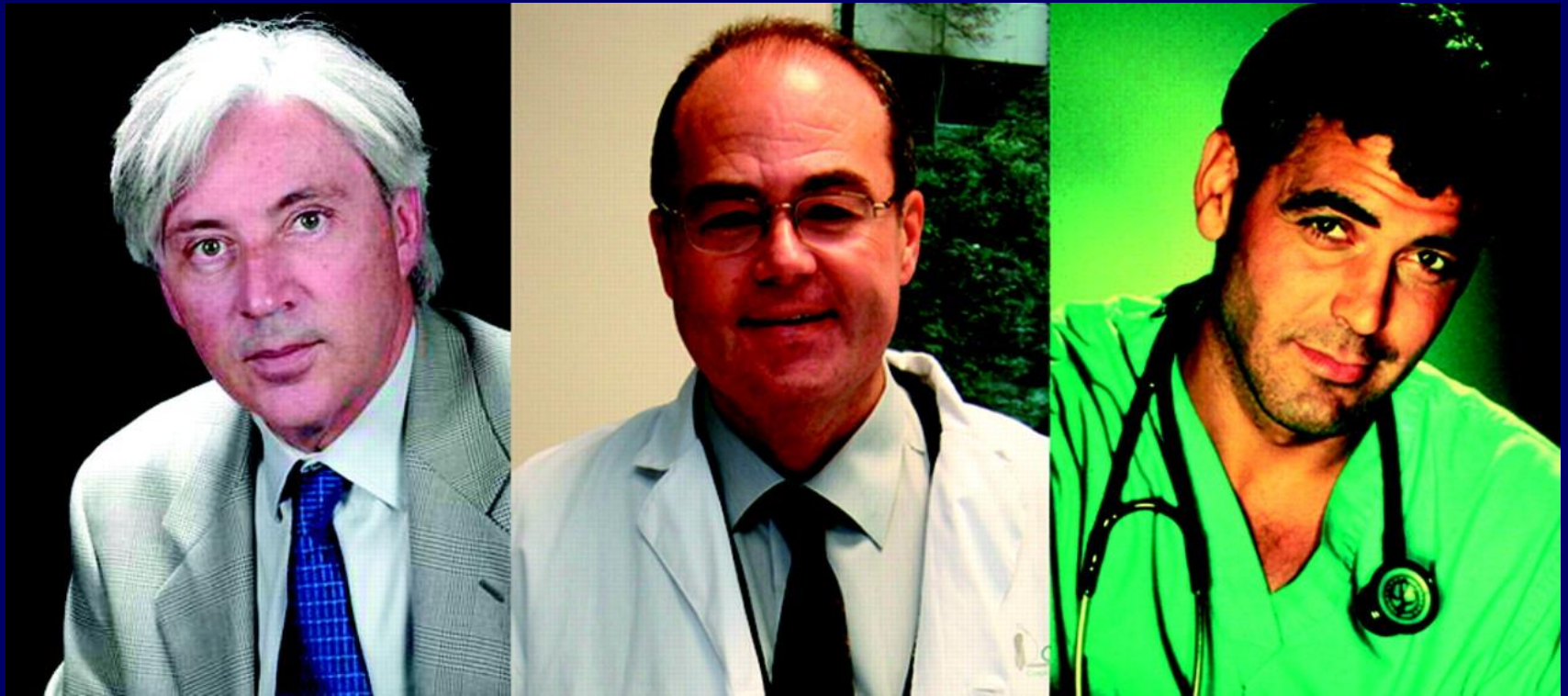
Główne wskaźniki oceny

Wzrost (w cm) i oceniany punktowo
„wskaźnik dobrego wyglądu”

Wyniki

- Chirurdzy są znamienne wyżsi od internistów (średni wzrost 179,4 vs 172,6 cm; $P=0,01$)
- Osoby w grupie kontrolnej mają znamienne wyższy „wskaźnik dobrego wyglądu” w porównaniu z chirurgami (średnia 5,96 vs 4,39; $P=0.013$) i internistami (5,96 vs 3,65; $P=0,003$)
- Chirurdzy mają znamienne wyższy „wskaźnik dobrego wyglądu” niż interniści (4,39 vs 3,65; 95% zakres ufności 0,25 do 1,23; $P=0,010$)

Przykłady



Chirurg

Internista

Grupa kontrolna
(George Clooney)

Wnioski

Chirurdzy są wyżsi i przystojniejsi niż interniści, ale aktorzy grający role lekarzy są przystojniejsi od obu tych grup. Nie wiadomo, czy te fenotypowe różnice mają podłoże genetyczne, czy środowiskowe.