

# *Meta-analiza danych*

Krzysztof Misztal

**Uniwersytet Jagielloński**  
Wydział Matematyki i Informatyki  
Informatyka

09.11.2009

- 1 Co to jest meta-analiza danych?
- 2 Fixed effects method
- 3 Heterogeniczność
- 4 Random effects method
- 5 Ograniczenia meta-analzy
- 6 Bibliografia

# The BIG idea

More studies  
are  
better than  
less studies.

# The BIG idea

Niektóre badania mogą być *ważniejsze* od pozostałych, więc dobre by było, aby istniała metoda umożliwiająca sprawdzenie zbioru badań - *to właśnie meta-analiza.*

# Co to jest meta-analiza danych?

- Ilościowy przegląd literatury.
- Synteza wyników poprzednich badań.
- Analiza próbki składającej się z badań, a nie z osób fizycznych.

# Dzięki meta-analizie jesteśmy w stanie odpowiedzieć na pytania

## Jak duży jest wpływ:

- treningu wytrzymałościowego na ciśnienie spoczynkowe krwi?
- wpływu używania kreatyny na wyniki w sprincie?
- i wiele innych...

## Dlaczego meta-analiza jest ważna?

Naukowcy zwykli myśleć, że celem pojedynczego badania jest podjęcie decyzji, czy dany efekt jest słuszny statystycznie.

Ale nie warto pokładać zbyt dużej wiary w pojedyncze badania, bo czasem może się ono okazać nieistotne statystycznie lub źle przeprowadzone.

Gdy przeprowadzono kilka badań na temat tego samego efektu, zaczęto się zastanawiać jak zaczęto się zastanawiać jak zebrać tę wiedzę i wyznaczyć rzeczywisty wpływ danego efektu.

Niech  $\theta$  będzie efektem leczenia lub stosowania środka zaradczego.

Niech  $\hat{\theta}_i$  będzie estymatorem efektu w  $i$ -tym badaniu oraz niech  $\hat{\sigma}_i^2$  będzie jego wariancją.

Jeśli znamy  $\sigma_i^2$  wtedy najlepszym estymatorem  $\theta$  jest

$$\hat{\theta} = \frac{\sum_i w_i \hat{\theta}_i}{\sum_i w_i},$$

gdzie  $w_i = 1/\sigma_i^2$ .



Jeśli  $\sigma_i^2$  nie jest znana to stosujemy  $w_i = \frac{1}{\hat{\sigma}_i^2}$ .

Wzór z wagami zakłada, że wartość  $\theta$  jest stała (fixed) w każdym badaniu, a różnice wynikają tylko z błędów pomiarów. Ten właśnie model nazywamy **fixed effect model**.

# Jednorodność badań

Istnieje wiele sposobów badania heterogeniczności:

- graficzne: forest plot, Galbraith plot...
- *numeryczne*: najczęściej używanym jest Q-test Cochran'a

$$Q = \sum w_i (\hat{\theta}_i - \hat{\theta})^2$$

gdzie:  $\hat{\theta}_i$  - jest estymatorem  $\theta_i$ , natomiast  $\hat{\theta}$  jest estymatorem  $\theta$ .

# Random effects method

Zakładamy, że wartość  $\theta_i$  w  $i$ -tym badaniu została wybrana losowo z rozkładu normalnego o średniej równej  $\theta$ .

Dokładniej dla  $x_i$  - obserwowanego efektu z  $i$ -tego badania mamy:

$$x_i = \theta_i + e_i, E(e_i) = 0, \text{Var}(e_i) = \sigma_i^2$$

$$\theta_i = \theta + u, E(u) = 0, \text{Var}(u) = \tau^2$$

- $\sigma_i^2$  - to wariancją poszczególnego badania
- $\tau^2$  - to wariancja między badaniami
- $\text{Var}(x_i) = \sigma_i^2 + \tau^2$
- $w_i = \frac{1}{\sigma_i^2 + \tau^2}$

**Zastosowanie tego wymaga estymowania  $\sigma^2$   
oraz  $\tau^2$ .**

$\sigma_i^2$  jest estymowana tak jak w *fixed effects method*

$\tau^2$  - najbardziej znanym jest estymator DerSimonian'a i Laird'a

$$\tau^2 = \begin{cases} \frac{Q-(k-1)}{\sum w_i - \frac{\sum w_i^2}{\sum w_i}} & Q \geq k - 1 \\ 0 & \text{wpp} \end{cases}$$

# Ograniczenia meta-analazy

- Skupia się na średni efekcie danego czynnika oraz różnicach pomiędzy poszczególnymi badaniami  
**A najbardziej się liczy wpływ/skutki danego czynnika dla jednostki**
- Porównywane są tylko czynniki publikowane:  
problemy z bazami danych, re-publikacje,  
problemy językowe, itd.

# Bibliografia

Darmowy program do wykonywania meta-analiz

<http://www.mix-for-meta-analysis.info/>

<http://www.mp.pl/artykuly/?aid=10529>