

Witold Simon

**META-ANALIZA W BADANIACH NAD SKUTECZNOŚCIĄ PSYCHOTERAPII.  
CZĘŚĆ I: PYTANIA BADAWCZE, PRZEGLĄD LITERATURY, KODOWANIE DANYCH.  
META-ANALYSIS IN PSYCHOTHERAPY OUTCOME RESEARCH.  
PART I: RESEARCH QUESTIONS, LITERATURE REVIEW, DATA CODING.**

Klinika Nerwic Instytutu Psychiatrii i Neurologii, Warszawa.

Kierownik: prof. dr hab. med. Maria M. Siwiak-Kobayashi.

**Streszczenie**

Jest to pierwszy artykuł z cyklu poświęconego meta-analizie w badaniach nad skutecznością psychoterapii. W tej części omówiono podstawowe założenia meta-analizy, w tym ograniczenia tej procedury. Opisano proces stawiania adekwatnych pytań badawczych. Zaprezentowano zasady kierujące, leżącym u podłoża meta-analizy, przeglądem literatury. Przedstawiono również sugestie dotyczące kodowania danych.

**Abstract**

This is the first part of the three-part series of papers devoted to the presentation of meta-analysis in psychotherapy outcome research. In this section, the basic meta-analytical assumptions, are outlined, including limitations of this procedure. Specifically, the validity of meta-analytical research questions is explained. The rules of performing an accurate literature review are presented. Moreover, suggestions for precise coding are listed.

Słowa kluczowe: meta-analiza, psychoterapia, pytania badawcze, przegląd literatury, kodowanie danych

Key words: meta-analysis, psychotherapy, research questions, literature review, data coding

**Wstęp**

Praca niniejsza stanowi pierwszą część cyklu poświęconego badaniom meta-analitycznym skuteczności psychoterapii. Meta-analiza uważana jest przeważnie za procedurę precyzyjnie oszacowującą rezultaty poszczególnych badań a poprzez to umożliwiającą rzetelną weryfikację zasobów literatury [1,2,3]. Jednocześnie jednak jej rozbudowana i skomplikowana metodologia jest często krytykowana [4,5,6,7]. Celem niniejszego cyklu jest więc omówienie zarówno głównych aspektów metodologicznych meta-analizy, jak i przedstawienie korzyści oraz ograniczeń płynących ze stosowania tej procedury badawczej. Ponieważ charakter niniejszej pracy umożliwia jedynie naszkicowanie powyższych zagadnień, zachęca się czytelnika do zapoznania się z pozycjami książkowymi [1,2,3] oraz zasobami *on-line* [8], wyczerpująco opisującymi zjawisko meta-analizy.

Psychoterapia, która stanowi osnowę niniejszych rozważań poświęconych meta-analizie, pomimo tego, że uchodzi za skuteczną metodę leczenia [9,10], to jednak trudno poddaje się ocenie badaczy, z racji na wieloczynnikowe uwarunkowanie procesu zmiany ludzkiego funkcjonowania [11]. W efekcie czytelnik zainteresowany oceną skuteczności psychoterapii ma do czynienia z zbyt różnorodnie zdefiniowanymi konstruktami oraz nadmiernie rozbieżnymi narzędziami badawczymi [12]. Również inne braki metodologiczne, jak chociażby niedostateczna reprezentatywność

poszczególnych grup pacjentów, czy częste pomijanie zmiennych związanych z terapeutą [5], przyczyniają się do utrzymujących się niejasności co do skuteczności tej formy leczenia.

Badacze, także ci zajmujący się psychoterapią, poszukują więc metod umożliwiających rzetelne opisanie zbioru danych pochodzących z wielu badań. Historycznie pierwszą taką próbę, w tym wypadku dotyczącą skuteczności szczepionki duru brzuszego, podjął Pearson w 1904 roku [13]. Dla każdego rozpatrywanego badania korelował on najpierw fakt podania szczepionki z wystąpieniem odporności lub zgonem. Następnie obliczał średnią arytmetyczną korelacji, uzyskując w ten sposób ocenę średniej skuteczności działania szczepionki. Współczesna era meta-analizy zaczęła się od badań Glassa [14], który wprowadził termin meta-analiza na określenie procedur statystycznych umożliwiających podsumowanie wyników pochodzących z niezależnych od siebie badań. Od tego czasu techniki meta-analityczne są szeroko stosowane, szczególnie w naukach behawioralnych, medycznych i fizyce, a ich liczba systematycznie rośnie (Tabela 1.)

**Tabela 1.** Liczba opublikowanych, na przestrzeni lat 1970-2009, meta-analiz zarejestrowana w internetowych bazach *Medline*, *PsycInfo*, *PsycBooks* oraz *PsycArticles*.

Lata	Liczba wszystkich meta-analiz	Liczba meta-analiz odnoszących się do psychoterapii
1970-1979	18	11
1980-1989	1640	110
1990-1999	9011	251
2000-2009	26367	668

Źródło: opracowanie własne. Stan na: 30 czerwca 2009 roku

Następujące cechy meta-analizy zadecydowały o upowszechnieniu tej procedury badawczej: 1) integracja i interpretacja zbiorczych wyników różnorodnych badań [2]; 2) oparcie analizy całość literatury zajmującej się danym problemem, zamiast polegania na pojedynczych doniesieniach [2]; 3) precyzyjne szacowanie podobieństw i różnic w metodologii poszczególnych prac [2]; 4) statystyczna, nie subiektywna ocena wielkości efektu [15]; 5) rzetelne opracowanie wniosków, w oparciu o zintegrowane wyniki, co ułatwia opracowywanie wytycznych dla terapeutów, pacjentów oraz polityków [1].

Meta-analiza stanowi wartościową alternatywę dla przeglądu literatury, gdyż ten przeważnie cechuje się: 1) tendencyjnością publikacyjną specjalistycznych czasopism nie uwzględniających badań pozostających poza kręgiem ich zainteresowań; 2) stronniczością autora objawiającą się zbyt subiektywnymi kryteriami selekcyjnymi lub wybiórczą oceną wyników; 3) znaczną rozbieżnością jakości metodologicznej włączonych badań; 4) brakiem możliwości analizy zmiennych moderujących.

Wczesne prace meta-analityczne [16] spotykały się z intensywną krytyką, czasami o zaskakująco emocjonalnym charakterze. Np. Eysenck używał określeń typu „mega-głupota” (*mega-silliness*) na włączanie do meta-analizy prac słabych metodologicznie [4]. Wytykał Glassowi, że włączył do swej meta-analizy jeden ze jego artykułów, o którym on sam „nie ma pojęcia” oraz zapowiadał, że badania meta-analityczne „zaprowadzą badaczy do ciemnego wieku psychologii” [4]. Na poparcie tych uwag, Eysenck nie przedstawił jednak wystarczająco rzeczowych argumentów.

Współczesne obszary krytyki meta-analizy dotyczą przeważnie [5,17]: 1) zbytniego polegania na statystyce z pominięciem intuicji; 2) włączania do jednej meta-analizy badań zbyt rozbieżnych metodologicznie; 3) braku rutynowego ważenia wielkości efektu; 4) niewystarczająco precyzyjnego definiowania konstruktów, co utrudnia operacjonalizację wyniku obserwowanego w niezależnych od siebie badaniach; 5) braku konsensusu odnośnie hipotetycznych mediatorów i moderatorów; 6) koncentrowania się na badaniach nad anglosaską populacją; 7) pomijania nieopublikowanych prac; 8)

dużej liczby rodzajów wielkości efektu; 9) włączania badań słabych metodologicznie.

Szczególnie ostatnie zastrzeżenie, związane z wprowadzeniem do meta-analizy niewłaściwych danych, wskazuje na zagrożenie uzyskania niewartościowych wyników (*garbage in and garbage out*) [17]. W dalszej części pracy odniesiono się do powyższych krytycznych kwestii.

### Metodologia meta-analizy - podstawowe etapy

Podstawowe etapy metodologiczne meta-analizy [1,18,19] zarysowano w Tabeli 2. a bardziej szczegółowo opisano w dwóch kolejnych artykułach niniejszego cyklu.

**Tabela 2.** Podstawowe etapy metodologiczne w meta-analizie

Etap metodologiczny	Istota danego etapu	Cele szczegółowe	Część cyklu
Pytania badawcze	<ul style="list-style-type: none"> <li>➔ formułowanie pytań <i>a priori</i></li> <li>➔ precyzyjna operacjonalizacja konstruktów badawczych</li> </ul>	<ul style="list-style-type: none"> <li>✓ uwzględnianie : a) rodzaju i wielkości grupy badanej, b) typu interwencji terapeutycznej, c) rodzaju grupy porównawczej, d) wyniku leczenia, e) wielkości efektu, f) narzędzi badawczych, g) metod statystycznych,</li> <li>✓ koncentrowanie się na badaniach podejmujących szczegółowe kwestie terapeutyczne,</li> <li>✓ zdefiniowanie kryteriów włączania i niewłączania badań do meta-analizy,</li> <li>✓ stosowanie skali kategoryzującej badania według wartości metodologicznej.</li> </ul>	I.
Przegląd literatury	<ul style="list-style-type: none"> <li>➔ dotarcie do jak największej liczby badań</li> <li>➔ określanie i oszacowanie tendencyjności publikacyjnej</li> </ul>	<ul style="list-style-type: none"> <li>✓ stosowanie synonimów i wyrazów bliskoznacznych,</li> <li>✓ korzystanie z elektronicznych baz danych z artykułami, rozdziałami, książkami oraz nieopublikowanymi pracami doktorskimi i magisterskimi,</li> <li>✓ prowadzenie ręcznego przeglądu spisu literatury,</li> <li>✓ docieranie do streszczeń konferencyjnych,</li> <li>✓ kontaktowanie się z kluczowymi badaczami z danej dziedziny,</li> <li>✓ określenia przyczyn potencjalnej tendencyjności publikacyjnej,</li> <li>✓ wybór metody szacującej tendencyjność publikacyjną.</li> </ul>	I.
Kodowanie danych	<ul style="list-style-type: none"> <li>➔ rzetelne kodowanie danych z poszczególnych meta-analizowanych prac</li> </ul>	<ul style="list-style-type: none"> <li>✓ określenie jasnej i precyzyjnej zmiennych oraz procedury kodowania,</li> <li>✓ osobne kodowanie badań „mocnych” metodologicznie a osobne „słabych”,</li> <li>✓ kodowanie przez dwa niezależne dwuosobowe zespoły koderów,</li> <li>✓ przeszkolenie i monitorowanie pracy koderów,</li> <li>✓ ocenienie stopnia rzetelności między koderami - statystyka kappa.</li> </ul>	I.
Analiza danych	<ul style="list-style-type: none"> <li>➔ wybór i obliczanie wielkość efektu</li> <li>➔ obliczanie sukcesu terapeutycznego</li> </ul>	<ul style="list-style-type: none"> <li>✓ wybór wielkości efektu w zależności od: a) procedur statystycznych użytych w analizowanych badaniach; b) rodzaju zmiennych,</li> <li>✓ kategoryzowanie wielkości efektu,</li> <li>✓ przeliczanie wartości między poszczególnymi wielkościami efektu,</li> <li>✓ przeliczanie wielkości efektu na centyle i procenty,</li> <li>✓ stosowanie binominalnej wielkość efektu,</li> </ul>	II.

Etap metodologiczny	Istota danego etapu	Cele szczegółowe	Część cyklu
	→ testowanie homogeniczności	<ul style="list-style-type: none"> <li>✓ opisanie przyczyn heterogeniczności we włączonych badaniach,</li> <li>✓ kliniczna vs. statystyczna ocena homogeniczności,</li> <li>✓ wybór testu homogeniczności,</li> <li>✓ testowanie hipotezy homogeniczności,</li> <li>✓ kategoryzowanie procentu heterogeniczności</li> </ul>	II.
	→ poszukiwanie moderatorów i mediatorów	<ul style="list-style-type: none"> <li>✓ opisanie i testowanie zmiennych moderujących,</li> <li>✓ opisanie i testowanie zmiennych mediujących.</li> </ul>	II
	→ modele wielkości efektu	<ul style="list-style-type: none"> <li>✓ korzystanie ze stałego modelu wielkości efektu – w przypadku stwierdzonej homogeniczności,</li> <li>✓ korzystanie z losowego modelu wielkości efektu – w przypadku stwierdzonej heterogeniczności,</li> <li>✓ korzystanie z modelu hybrydowego – w niejednoznacznych przypadkach,</li> <li>✓ ważenie wielkości efektu</li> </ul>	III.
	→ przycinanie danych	<ul style="list-style-type: none"> <li>✓ określenie wyników nietypowo odstających od pozostałych rezultatów,</li> <li>✓ eliminacja odstających danych</li> </ul>	III.
	→ ocena replikacji meta-analizy	<ul style="list-style-type: none"> <li>✓ porównanie wielkości efektu,</li> <li>✓ porównanie istotności statystycznych.</li> </ul>	III.
Wnioski	<ul style="list-style-type: none"> <li>→ wyciąganie wniosków klinicznie znaczących</li> <li>→ wyciąganie wniosków metodologicznie istotnych</li> </ul>	<ul style="list-style-type: none"> <li>✓ wyciąganie wniosków w oparciu o uzyskaną ocenę skali obserwowanych wyników,</li> <li>✓ ostrożne wnioskowanie w przypadku: a) prac bazujących na małej liczbie publikacji, b) prac cechujących się znaczną heterogenicznością.</li> </ul>	III.

Źródło: opracowanie własne.

### Formułowanie pytań badawczych

Jakość metodologiczna meta-analizy zależy od tego, na ile porównywane badania odnoszą się do zbliżonego konstruktów mierzonego przez standardowe narzędzia badawcze i analizowanego przy użyciu takich samych procedur statystycznych.

Pytania badawcze należy więc sformułować *a priori*, koniecznie z uwzględnieniem co najmniej: 1) rodzaju i wielkości grupy badanej oraz porównawczej; 2) opisu szkoły psychoterapeutycznej; 3) typu interwencji terapeutycznej; 4) operacjonalizacji wyniku leczenia; 5) typu wielkości efektu; 6) rodzaju narzędzi badawczych; 7) typu metod statystycznych. Wszystkie konstrukty powinny być jednoznacznie i szczegółowo zdefiniowane. Chociażby dlatego, że pacjenci są często leczeni nie do końca zgodnie z deklarowaną przez autora pracy orientacją psychoterapeutyczną [20]. Przy określaniu zmiennych zależnej i niezależnych zalecane jest koncentrowanie się na badaniach podejmujących szczegółowe kwestie, np.: ocenę skuteczności specyficznej psychoterapii lub nawet konkretnej techniki terapeutycznej wobec konkretnego objawu [21]. Pojawiają się również sugestie, że należy uszczegóławiać meta-analityczne pytania badawcze według stopnia klinicznej istotności poszczególnych prac, określanej w zależności od procentu pacjentów, którzy wyzdrowieli, poprawili się, nie zmienili się lub pogorszyli się podczas psychoterapii [22].

Na tym etapie należy jasno zdefiniować kryteria dotyczące włączania lub niewłączania badań

do meta-analizy. Jednocześnie, metodologia badań włączanych do meta-analizy powinna podlegać rygorystycznej ocenie, poczynając od rozróżnienia prac randomizowanych od nierandomizowanych. Rozróżnieniu takiemu służyć może skala oceny wartości metodologicznej danego badania [23]. Zakłada ona następujące punkty odcięcia: 0.1 – badanie bez grupy kontrolnej, 0.2 – badanie z nielosowo dobraną grupą kontrolną, 0.9 – badanie z losowo dobraną grupą kontrolną nie poddawaną leczeniu, 1.0 – badanie z losowo dobraną grupą porównawczą opartą o terapię placebo.

Kwestia czy w ogóle należy włączać do meta-analizy badania bez grupy porównawczej lub też bez losowego doboru pacjentów do grup badanych pozostaje wciąż dyskusyjna [24,25]. W dużej mierze zależy ona od tego, czy meta-analityk przywiązuje większą wagę do klinicznej czy też statystycznej reprezentatywności, a co za tym idzie, czy pragnie przedstawić wynik meta-analizy w świetle skuteczności (*efficacy*) czy też skuteczności (*effectiveness*) [10]. Nie budzi natomiast kontrowersji zalecenie by nie włączać do meta-analizy różnych prac opartych o tą samą grupę badaną, gdyż groziłoby to uzyskaniem nieadekwatnej wielkości efektu.

### **Przegląd literatury**

Meta-analizę należy poprzedzić szczegółowym przeglądem zasobów literatury, który ma pomóc odpowiedzieć na pytanie na ile dana dziedzina nauki dojrzała już do przeprowadzenia meta-analitycznych badań. Przeprowadzenie dogłębnego przeglądu zakłada posługiwanie się synonimami i wyrazami bliskoznacznymi słów kluczowych, co zwiększa szansę na dotarcie do większej liczby badań. W przypadku terminu psychoterapia (*psychotherapy*) pomocne może być posługiwanie się też pojęciami: *therapy, treatment, counseling, intervention, help, guidance, training, cure, care, rehabilitation, healing, analysis, remedy*, itp. Przegląd literatury powinien opierać się nie tylko na elektronicznych bazach danych zawierających artykuły (np. *Medline* czy *PsycInfo*), ale także na bazach z książkami (np. *PsycBook*). Ręczny przegląd bibliografii z dostępnych prac, także zwiększa szansę na poszerzenie zbioru meta-analizowanych badań. Pomocne może być również skontaktowanie się z kluczowymi badaczami z danej dziedziny, którzy mogą posiadać informacje o pracach, które z różnych powodów nie znalazły się w elektronicznych bazach danych. Warto również dotrzeć do nieopublikowanych prac doktorskich i magisterskich (np. w *Dissertation Abstracts International*), jak i do streszczeń doniesień konferencyjnych.

Nieopublikowane prace tworzą problem badań napisanych do szuflady (*file drawer problem*), czasami określane jako problem tzw. szarej literatury (*grey literature*). Generalnie, badania nie są publikowane dlatego, że: 1) są słabe metodologicznie; 2) niosą ze sobą nieistotne statystycznie wyniki; 3) znajdują się w przedłużającej się recenzji; 4) autorom zabrakło wytrwałości w dopracowaniu materiału. Pomimo tego, że zakłada się, że szara literatura jest mniej wartościowa metodologicznie [26] niż prace pochodzące z czasopism recenzowanych [27], to dowiedziono, że wykluczenie jej z meta-analizy, może istotnie zawyżać wartość wielkości efektu, nawet o 15% [27]. Niemniej, tylko jedna trzecia meta-analiz [27] uwzględnia szarą literaturę, która gdy jednak jest już włączona do meta-analizy to stanowi od 4% do 75% badanych prac. Z kolei, weryfikując streszczenia konferencyjne, należy pamiętać, że tylko część z nich ukazuje się później w formie artykułów, co zajmuje przeważnie od 2 do 8 lat [28]. Wykluczenie ich jednak z meta-analizy zawyża wielkość efektu średnio o 38% [28].

Problem nieopublikowanych prac jest jedną z głównym przyczyn tendencyjności publikacyjnej (*publication bias*). Tendencyjność ta bywa również spowodowana pomijaniem publikacji wydanych w innych językach niż angielski, co określa się mianem tendencyjności wieży Babel (*The Tower of Babel bias*) [29]. Niektórzy uważają wprawdzie, że zawężenie poszukiwań do języka angielskiego nie prowadzi do błędów w obliczaniu wielkości efektu [30], wydaje się jednak, że meta-analityk powinien, w miarę możliwości, poszerzyć pole poszukiwań o inne języki kongresowe oraz język rodzimy. Tendencyjność publikacyjna może też wynikać z włączania do badań pacjentów bardziej zaburzonych, czyli potencjalnie w większym stopniu podatnych na leczenie, jak i na jego efekty uboczne. W takim

wypadku zaleca się meta-analizę warstwową (*stratified meta-analysis*) [6] różnicującą wyniki w zależności od stopnia ryzyka związanego z leczeniem osób bardziej zaburzonych.

Liczba nieopublikowanych prac przeważnie pozostaje jednak nieznaną, jako że meta-analityk, nawet pomimo dołożenia wszelkich starań, nie jest w stanie dotrzeć do wszystkich źródeł. W związku z tym należy precyzyjnie oszacować skalę tendencyjności publikacyjnej. Można tego dokonać na kilka sposobów.

Pomocne jest np. oszacowanie liczby badań napisanych do szuflady, które choć nie dostępne mogłyby zwiększyć prawdopodobieństwo błędu I rodzaju odnośnie hipotezy zerowej i poprzez to wpłynąć na poziom istotności całościowych wyników meta-analizy. Innymi słowy, należy określić liczbę badań, których ewentualne dodanie do meta-analizy może spowodować przesunięcia pierwotnej istotności meta-analizy z poziomu np. ,05, 0,01 lub ,001 do poziomu  $p > ,05$ , czyli do granicy nieistotności statystycznej. W tym celu można się posłużyć Tabelą 3. [31]. Przykładowo, jeżeli poddano meta-analizie 15 badań o średniej istotności 0,05, to wspomniana tabela wskazuje, że liczba niedostępnych badań, które mogą potencjalnie przesunąć istotność powyżej 0,05 wynosi w tym przypadku 210. Generalnie, za odporne wobec problemu niedostępnej szarej literatury uznaje się takie badanie, którego oporność nie przekroczy poziomu  $5k + 10$ , gdzie  $k$  oznacza liczbę dostępnych badań [30]. W podanym przykładzie meta-analizy opartej na 15 badaniach oznacza to, że istotność na poziomie 0,05 będzie utrzymana o ile liczba nieznanymi nieopublikowanych badań nie przekroczy 85 ( $5 \times 15 + 10 = 85$ ). A jako że  $85 < 210$ , można więc uznać, że dana meta-analiza jest odporna na zagrożenie problemem szarej literatury.

**Tabela 3.** Tolerancja dla zerowych wyników będących funkcją pierwotnej średniej istotności danej meta-analizy oraz liczby włączonych badań.

Liczba włączonych badań	pierwotna średnia istotność		
	,05	,01	,001
1	1	2	4
2	4	8	15
3	9	18	32
4	16	32	57
5	25	50	89
6	36	72	128
7	49	98	173
8	64	128	226
9	81	162	286
10	100	200	353
15	225	450	795
20	400	800	1412
25	625	1250	2206
30	900	1800	3177
40	1600	3200	5648
50	2500	5000	8824

Liczba badań w tabeli odnosi się do liczby wszystkich badań (starych i nowych) wymaganej do przesunięcia pierwotnej istotności z poziomu ,05, 0,01 lub ,001 do poziomu  $p > ,05$  (tzn. do granicy nieistotności statystycznej). Źródło: zaadaptowano z [31].

Innym sposobem oszacowania skali tendencyjności publikacyjnej jest metoda „pełnej szuflady” („*file drawer*” method) [32]. Metoda ta bazuje na założeniu, że wynik meta-analizy osiąga istotne wartości dla  $Z > Z_{1-\alpha/2}$  w oparciu o test oparty na połączonych standardowych ocenach typu  $Z$  pochodzących z opublikowanych badań:

$$Z = \left( \sum_{i=1}^k Z_i \right) / \sqrt{k} ,$$

gdzie:  $Z_i$  – standardowa ocena typu  $Z$ ,  $k$  – liczba włączonych badań. Standardowa ocena typu  $Z$  odzwierciedla odległość, podawaną w postaci odchyleń standardowych, o którą pierwotna ocena oddalona jest od średniej.

Metoda ta umożliwiła określenie liczby nieopublikowanych badań ze średnim obserwowalnym efektem na poziomie zero, które należałoby dodać tak by wynik  $Z$  nie był już istotny. Liczbę dodatkowych badań potrzebnych do uczynienia meta-analizy nieistotną, oblicza się według wzoru [33]:

$$k_o > -k + \left( \sum_{i=1}^k Z_i \right)^2 / (Z_{1-\alpha/2})^2 ,$$

gdzie:  $k_o$  – liczba dodatkowych badań potrzebnych do uczynienia meta-analizy nieistotną,  $Z_i$  – standardowa ocena typu  $Z$ ,  $k$  – liczba włączonych badań.

Jeżeli  $k_o$  jest wystarczająco duże, wtedy jest mało prawdopodobne, że istnieje wiele nieopublikowanych badań mogących wpłynąć znacząco na istotność obserwowanego efektu.

Kolejną metodę szacowania tendencyjności publikacyjnej proponują Hedges i Olkin [34]. Sugerują oni liczenie zarówno pozytywnych (istotnych statystycznie) jak i negatywnych (nieistotnych statystycznie) wyników badań które zostały włączone do meta-analizy. Proponują oni wzór:

$$p = p^+ / (p^+ + p^-),$$

gdzie:  $p^+$  - liczba istotnych statystycznie wyników,  $p^-$  - liczba nieistotnie statystycznych wyników.

#### **Kodowanie danych**

Kodowanie danych wymaga jasno określonej procedury i precyzji w jej przestrzeganiu. Szanse na większą rzetelność rosną, jeżeli każda praca włączona do meta-analizy będzie kodowana przez dwa niezależne dwuosobowe zespoły koderów. Osoby te powinny otrzymać szczegółowy opis zmiennych (Tabela 4.), które będą kodowane [2] oraz przejść stosowne szkolenie, a wszystko to w celu zminimalizowania potencjalnych rozbieżności w kodowaniu poszczególnych danych. Lista potencjalnych rozbieżności jest długa i może dotyczyć np.: 1) rodzaju terapii (np. stopnia zgodności postępowania terapeutów z teoretycznym modelem terapii); 2) grupy badanej (np. reprezentatywności wobec klinicznej populacji); 3) liczby sesji (np. wliczania lub pomijania sesji kwalifikacyjnej i sesji katamnesticznych); 4) psychoterapeutów (np. ich formalnych kwalifikacji); 5) potencjalnego konfliktu

interesu badacza (np. faktu, że bada skuteczność terapii, z ośrodka w którym sam pracuje); 6) kwestii metodologicznych (np. rodzaju wielkości efektu i sposób jego ważenia); 7) danych publikacyjnych (np. oceny jakości czasopisma), itp. Generalnie, należy osobno kodować wyniki badań „mocnych” metodologicznie a osobno „słabych”. Określenie punktu odcięcia w tym kontinuum jest oczywiście arbitralne, ale powinno być wyraźnie określone.

**Tabela 4.** Przykładowa procedura kodowania danych

Zmienna	Wskazówki dla koderów
Dane bibliograficzne	✓ nazwisko autora i rok publikacji, np. Fonagy, 1998
Rodzaj publikacji	✓ 1. czasopismo, 2. książka lub rozdział, 3. praca doktorska lub magisterska, 4. streszczenie konferencyjne, 5. inne
Koder	✓ numer kodera: np.: 1
Opis grupy badanej	✓ średnia wieku, % kobiet, % rasy kaukaskiej, % innych grup rasowych, % singli, % rozwodników, % wdowców/wdów, % osób żyjących w separacji
Rodzaj zaburzenia (np. zaburzenia osobowości typ C)	✓ % pacjentów z zaburzeniami z klasteru (skupienia) C w całej grupie badanej ✓ % osobowość unikająca; % osobowość depresyjna; % osobowość obsesyjno-kompulsyjna; ✓ kryteria diagnostyczne (1. DSM-III, 2. DSM-III-R, 3. DSM-IV, 4. DSM-IV-R, 5. inne)
Opis terapii	✓ modalność (1. indywidualna, 2. grupowa, 3. grupowa i indywidualna, 4. inna); ✓ orientacja (1. psychodynamiczna, 2. psychoanaliza lub analiza, 3. poznawczo-behawioralna lub poznawcza, 4. dialektyczno-behawioralna, 5. eklektyczna lub integracyjna, 6. trening umiejętności społecznych, 7. inne); UWAGA: koduj więcej niż jedną kategorię, o ile to konieczne; ✓ miejsce prowadzenie terapii (1. stacjonarna, 2. dzienna, 3. ambulatoryjna, 4. mieszana, 5. inna); ✓ liczba sesji, UWAGA: koduj tylko rzeczywistą liczbę sesji, nie liczbę sesji planowanych; ✓ okres badania katamnesticznego (liczba miesięcy jaka upłynęła od końca terapii); ✓ opis badania katamnesticznego (1. z terapią, 2. bez terapii)
Projekt badawczy	✓ przypisanie do grupy badanej (1. losowe, 2. naturalistyczne, 3. inne); ✓ przypisanie do grupy kontrolnej/porównawczej (1. losowe, 2. naturalistyczne, 3. inne)
Grupa kontrolna, porównawcza	✓ rodzaj grupy kontrolne/porównawczej w okresie terapii (1. ista oczekujących, 2. brak jakichkolwiek oddziaływań terapeutycznych, 3. typowe leczenie, 4. minimalny kontakt, 5. placebo w formie uwagi, 6. farmakoterapia, 7. relaks, 8. porównywana forma terapii, 9. inna diagnoza niż badana, 10. inny typ grupy kontrolnej, 11. brak grupy kontrolnej); ✓ rodzaj grupy kontrolne/porównawczej w okresie katamnesticznym (kategorie takie same jak w okresie terapii);
Wielkość efektu	✓ rodzaj wielkości efektu (1. porównanie przed/po, 2. porównanie przed/po grupy badanej vs. grupy kontrolnej, 3. porównanie przed/katamneza, 4. porównanie po/katamneza, 5. porównanie katamneza/katamneza grupy badanej vs. grupy kontrolnej); ✓ wielkość efektu oparta na (1. średnich i odchyleniach standardowych, 2. wartości t-testu, 3. wartości F-ANOVA, 4. wartości chi-kwadrat, 5. dychotomicznej lub politochomicznej proporcjonalności, 6. częściowej korelacji, regresji, współczynnika ścieżkowym, 7. analizie kowariancji, 8. inne, 9. mieszane); ✓ lista zastosowanych narzędzi badawczych.
dane	✓ grupa badana: średnie wartości zmiennych, wielkość grupy, odchylenia standardowe, na początku, na końcu leczenia, oraz podczas badania katamnesticznego; ✓ grupa kontrolna/porównawcza: średnie wartości zmiennych, wielkość grupy, odchylenia standardowe, na początku, na końcu leczenia, oraz podczas badania katamnesticznego; ✓ wielkość efektu ( <i>d</i> ) końcowa wartość.
Inne	✓ miejsce na opisanie problematycznych kwestie
Niejasne dane – kodować zawsze jako: “999”.	

Źródło: [20]. Przedruk za zgodą Informa Healthcare oraz Taylor and Francis Group.

Po zakończeniu kodowania danych do meta-analizy należy ocenić stopień zgodności między koderami z wykorzystaniem statystyki kappa, według następującego wzoru [35]:



(obserwowana zgodność niepodlegająca przypadkowi) / (maksymalna możliwa zgodność niepodlegająca przypadkowi).

Interpretację wartości kappę przedstawiono w Tabeli 5.

**Tabela 5.** Jakościowa klasyfikacja wartości współczynnika kappę jako stopnia zgodności niepodlegającego przypadkowi

Wartość kappę	Stopień zgodności
0.80 - 1.00	niemal idealny
0.60.-0.79	znaczący
0.40-0.59	średni
0.20-0.39	dostateczny
0.01-0.19	słaby
0.00	brak

Źródło: opracowanie własne

#### Adres do korespondencji:

Witold Simon,  
Klinika Nerwic IPiN, ul. Sobieskiego 9,  
02-957 Warszawa,  
[wsimon@ipin.edu.pl](mailto:wsimon@ipin.edu.pl)

#### Piśmiennictwo:

1. Wolf FM. Meta-analysis: Quantitative methods for research synthesis. Newbury Park, CA: Sage Publication; 1986.
2. Cooper H, Hedges LV. (red.) The Handbook of research synthesis. New York: Russel Sage Foundation; 1994.
3. Lipsey MW, Wilson DB. Practical meta-analysis. Thousand Oaks, CA: Sage Publication; 2001.
4. Eysenck HJ. An exercise in mega-silliness. American Psychologist, 1978; 33 (5): 517.
5. Matt GE, Navarro AM. What meta-analyses have and have not taught us about psychotherapy effect: a review and future directions. Clinical Psychology Review, 1997; 17(1): 1-32.
6. Thompson SG, Higgins JPT. Can meta-analysis help target interventions at individuals most likely to benefit? Lancet, 2005; 365: 341-346
7. Strube MJ, Hartman DP. A critical appraisal of meta-analysis. British Journal of Clinical Psychology. 1982; 21: 129-139
8. Cochrane. Cochrane handbook for systematic reviews of interventions. Oxford, UK. Uzyskano 24 marca 2009 z <http://www.cochrane.org/resources/handbook/index.htm>
9. Lambert MJ. (red.). Bergin & Garfield's handbook of psychotherapy and behavior change, 5<sup>th</sup> edition. New York: John Wiley & Sons; 2004.
10. Roth A, Fonagy P. What works for whom? A critical review of psychotherapy research. New York: The Guilford Press; 2005.
11. Gerrig R, Zimbardo PG. Psychology and life. 7<sup>th</sup> edition. Boston, MA: Allyn and Bacon Publishers; 2007.
12. Hill CE, Lambert MJ. Methodological issues in studying psychotherapy processes and outcomes. W: Lambert MJ. (red.). Bergin & Garfield's handbook of psychotherapy and behavior change, 5<sup>th</sup> edition. New York: Wiley; 2004, 84-135.
13. Pearson K. Report on certain enteric fever inoculation statistics. British Medical Journal. 1904; 3: 1243-1246.
14. Glass GV. Primary, secondary and meta-analysis of research. Educational Researcher. 1976; 5: 3-8.
15. Sack HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. New England Journal of Medicine. 1987; 316: 450-455.
16. Smith LM, Glass GV. Meta-analysis of psychotherapy outcome studies. American Psychologist. 1977; 32 (9): 752-760.
17. Hunt M. How science takes stock. New York: Russell Sage Foundation; 1997.

18. Durlak JA, Lipsey MW. A practitioner's guide to meta-analysis. *American Journal of Community Psychology*. 1991; 19 (3): 291-332.
19. Rosenthal R, DiMatteo MR. Meta-analysis: Recent developments in quantitative methods for literature review. *Annual Review of Psychology*. 2001; 52: 59-82.
20. Simon W. Follow-up psychotherapy outcome of patients with dependent, avoidant and obsessive-compulsive personality disorders - a meta-analytic review. *International Journal of Psychiatry in Clinical Practice*. 2009; 13 (2): 1-12.
21. Sharpe D. On apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review*. 1997; 17: 881-901.
22. Lambert MJ, Hansen NB, Bauer S. Assessing the clinical significance of outcome results. W: Nezu AM, Maguth-Nezu C. (red.). *Evidence-based outcome research: A practical guide to conducting randomized controlled trials for psychological interventions*. New York: Oxford University Press; 2008, 359-378.
23. Rosenthal R. Quality-weighting of studies in meta-analytic research. *Psychotherapy Research*. 1991; 1(1): 25-28.
24. Lambert MJ, Ogles BM. The efficacy and effectiveness of psychotherapy. W: Lambert MJ. (red.). *Bergin & Garfield's handbook of psychotherapy and behavior change*, 5<sup>th</sup> edition. New York: John Wiley & Sons; 2004, 139-193.
25. Chambless DL, Hollon SD. Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*. 1998; 66 (1): 7-18.
26. Cook DJ, Guyatt GH, Ryan G, Clifton J, Buck-Ingham L, Willian A, McIlroy W, Oxman AD: Should unpublished data be included in meta-analyses? *Journal of American Medical Association*. 1993; 269: 2749-2753.
27. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influences estimates of intervention effectiveness reported in meta-analyses? *Lancet*. 2000; 356: 1228-1231.
28. Stern JM, Simes RJ. Publication bias: Evidence of delayed publication in a cohort study of clinical research projects. *British Medical Journal*. 1997; 315: 640-645.
29. Egger M, Smith GD. Meta-analysis: Bias in location and selection of studies. *British Medical Journal*. 1998; 316: 61-66.
30. Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: Empirical study. *International Journal of Epidemiology*. 2002; 31: 115-123.
31. Rosenthal R. *Meta-analytic procedures for social research*. 2<sup>nd</sup> edition. Newbury, CA: Sage; 1991.
32. Rosenthal R. The "file drawer problem" and tolerance for null results. *Psychological Bulletin*. 1979; 86: 638-641.
33. Begg CB. Publication bias. In: Cooper H, Hedges LV. (editors) *The Handbook of research synthesis*. New York: Russel Sage Foundation; 1994
34. Hedges LV, Olkin I. Vote-counting methods in research synthesis. *Psychological Bulletin*. 1980; 88: 359-369.
35. McGinn T, Wyer PC, Newman TB, Keitz S, Leipzig R, Guyatt G. Measures of observer variability (kappa statistic). *Canadian Medical Association Journal*. 2004; 11 (171): 1369-1373.