

СОЦІОЛОГІЧНА ОСВІТА

УДК 303. 447.22

СЕРГІЙ ДЕМБІЦЬКИЙ,

кандидат соціологічних наук, молодший науковий співробітник відділу методології і методів соціології Інституту соціології НАН України

Оцінювання гетерогенності в метааналізі (на прикладі даних крос-національних досліджень)

Анотація

Статтю присвячено основам аналізу гетерогенності в рамках метааналізу. Описано суть оцінювання гетерогенності, а також розглянуто такі питання: статистична перевірка за допомогою Q-тесту; оцінювання істинної дисперсії величин ефектів за допомогою T^2 ; оцінювання пропорції спостережуваної мінливості, що вказує на істинну мінливість розмірів ефектів за допомогою I^2 ; розрахунок довірчих і передбачувальних інтервалів.

На основі практичного прикладу із застосуванням відношення шансів як величин ефектів демонструється розрахунок ключових показників гетерогенності (як емпіричну базу використано результати четвертої хвилі Європейського соціального дослідження).

Ключові слова: *метааналіз, величина ефекту, модель випадкових ефектів, показники гетерогенності*

Вступ

У своїй нещодавній публікації я розглянув ключові поняття й основи обчислень у метааналізі [Дембицкий, 2012]¹. При цьому особливу увагу було приділено двом моделям метааналізу — фіксованих та випадкових

¹ Тут мушу зазначити наявність чотирьох прикрих помилок, що їх було припущено у попередній статті. Їх розглянуто в Додатку. Помилки не позначилися на практичному прикладі — всі розрахунки здійснено правильно.

ефектів. У рамках моделі випадкових ефектів важливим завданням є оцінювання гетерогенності, що дає змогу: 1) проаналізувати, яка частка мінливості у величинах ефектів зумовлена відмінностями генеральних сукупностей, з яких вони отримані; 2) зробити прогноз щодо величини інтервалу, в який потрапляє більшість величин ефектів, безвідносно до того, з якої генеральної сукупності вони отримані; 3) сформулювати гіпотези про чинники мінливості, тобто про ті релевантні чинники, за якими генеральні сукупності відрізняються. Останнє уможлиблює роздільний аналіз контрастних підгруп величин ефектів і, відповідно, робить висновки дослідження гнучкішими завдяки прив'язці до контекстуальних особливостей, важливих з точки зору досліджуваного взаємозв'язку.

Слід наголосити, що далі йтиметься саме про оцінювання гетерогенності, а не її перевірку. Часто рішення стосовно гетерогенності/гомогенності результатів приймають у результаті статистичної перевірки (статистичний підхід). Проте в рамках даної статті я відштовхуюся від позиції Майкла Борнштейна з його співавторами [Borenstein et al., 2009], котрі пов'язують рішення щодо наявності гетерогенності результатів або її відсутності з природою аналізованих даних (концептуальний підхід). За статистичного підходу заведено розпочинати із використання моделі фіксованих ефектів, потім перевіряти її результати на гетерогенність і, в разі її підтвердження, братися до моделі випадкових ефектів. Натомість за концептуального підходу саме на підставі аргументів про природу емпіричного матеріалу (про те, чи слід вважати, що величини ефектів отримано із різних генеральних сукупностей) дослідник обирає одну із двох моделей. І тільки у разі вибору моделі випадкових ефектів здійснюють оцінювання гетерогенності, бо вона передбачена від початку¹.

Основи оцінювання гетерогенності

Коли йдеться про оцінювання гетерогенності в рамках метааналізу, то мають на увазі аналіз істинної (на противагу випадковим помилкам) мінливості величин ефектів. Разом із тим спостережувана мінливість включає обидва компоненти — і випадкові помилки, і істинну мінливість. Механізм, що дає змогу розділити ці компоненти, такий: 1) обчислити загальну величину мінливості, що спостерігається між дослідженнями, відібраними для аналізу; 2) оцінити очікувану мінливість для випадку, коли всі величини ефектів належать до однієї генеральної сукупності (тобто для випадку, коли сукупна мінливість ґрунтується суто на випадкових помилках); 3) знайти різницю між загальною й очікуваною мінливістю (так звану надлишкову варіацію), котра й указуватиме на величину гетерогенності або істинної мінливості [Borenstein et al., 2009: p. 108].

Статистична перевірка за допомогою Q-тесту. Традиційно оцінювання гетерогенності розпочинають з обчислення Q-статистики, що демонструє, чи відрізняється статистично значимо спостережувана мінливість від очікуваної, ґрунтованої на випадкових помилках.

¹ Докладніше про вибір моделі метааналізу в контексті оцінювання гетерогенності результатів див.: [Hunter, Schmidt, 2004: p. 393–399].

Q — показник, що має розподіл χ^2 , із кількістю ступенів свободи (df), що дорівнює $k - 1$, де k дорівнює кількості величин ефектів [Corcogan, Littel, 2010: р. 303]. Q є стандартизованим показником, тобто не залежить від метрики використовуваних величин ефектів. Базова формула для обчислення Q має вигляд:

$$1.1. Q = \sum_{i=1}^k W_i (Y_i - M)^2.$$

Еквівалентна формула, більш придатна для розрахунків:

$$1.2. Q = \sum_{i=1}^k W_i Y_i^2 - \frac{\left(\sum_{i=1}^k W_i Y_i \right)^2}{\sum_{i=1}^k W_i}.$$

Очікуване значення Q для випадку, коли всі величини ефектів належать до однієї генеральної сукупності, не відрізняється від df , оскільки є стандартизованою величиною. Отже, істинну варіацію у величинах ефектів можна розрахувати, віднявши від спостережуваного Q його очікуване значення, або df . Якщо ця величина негативна або несуттєво перевищує 0 (чи достатньою є величина відмінності між спостережуваним і очікуваним значеннями, визначається стандартним для використання критерію χ^2 чином — за допомогою p -значення), то перевірка не дала статистично значимих результатів. У протилежному разі приймається рішення про гетерогенність, що відкриває можливість подальшого аналізу.

Q -тест є перевіркою нульової гіпотези (згідно з якою всі дослідження описують той самий ефект або належать до однієї генеральної сукупності), що пов'язано з проблемами, характерними і для інших тестів статистичної значимості. По-перше, якщо статистично значимі результати вказують на гетерогенність, то незначимі ще не доводять її відсутності — можливе пояснення може приховуватися в малому обсязі вибірки чи у великих дисперсіях результатів окремих досліджень. По-друге, результати Q -тесту використовують тільки для перевірки нульової гіпотези і, відповідно, вони не є показником величини істинної мінливості. Це завдання виконують інші показники, про які йтиметься далі.

Оцінювання істинної дисперсії величин ефектів за допомогою T^2 . Показник T^2 повертає оцінювання мінливості від стандартизованої шкали до оригінальної метрики величин ефектів:

$$1.3. T^2 = \frac{Q - df}{C},$$

$$\text{де } C = \sum W_i - \frac{\sum W_i^2}{\sum W_i}.$$

Істинне значення дисперсії величин ефектів не може бути меншим за 0, тоді як T^2 може набувати негативних значень (при $Q < df$), що пов'язано із помилками вибірки. У такому разі T^2 приймають за 0. Якщо ж $Q > df$, тоді T^2 буде позитивним значенням, що базується на двох чинниках. Перший із

них — величина надлишкової мінливості ($Q - df$), другий — розмірність шкали, в якій обчислено величини ефектів.

Для оцінювання стандартного відхилення величин ефектів використовують T , яке можна знайти простим обчисленням квадратного кореня із T^2 . При визначенні T^2 не потрібне припущення про форму розподілу величин ефектів, але якщо є підстави вважати її нормальною, тоді T можна використовувати для знаходження інтервалу, в який потрапить задана пропорція всіх величин ефектів. Наприклад, 95% усіх можливих значень потрапить в інтервал, що дорівнює $1,96 \times T$.

Водночас такий спосіб побудови розподілу величин ефектів виправданий лише в тому випадку, коли і ці величини, і T оцінено правильно. Тому на практиці при побудові відповідних інтервалів для розподілу величин ефектів необхідно зважати на помилки в оцінюванні обох параметрів. Підстави відповідних розрахунків розглянуто далі.

Оцінювання пропорції спостережуваної мінливості, що вказує на істинну мінливість величин ефектів із використанням I^2 . I^2 є показником, похідним від Q :

$$1.4. I^2 = \left(\frac{Q - df}{Q} \right) \times 100\%.$$

Подібно до Q , показник I^2 є стандартизованим, тобто не залежить від оригінальної шкали величин ефектів і змінюється в діапазоні від 0% до 100%. Важливою відмітною рисою цієї статистики є також те, що вона не залежить від кількості аналізованих величин ефектів [Borenstein et al., 2009: р. 109–119].

При цьому слід пам'ятати — попри те, що I^2 дає змогу оцінити пропорцію істинної мінливості, він не дає жодної іншої інформації. Так, якщо відповідне значення близьке до 100%, це ще нічого не говорить про розподіл величин ефектів, котрі в цьому разі можуть потрапляти як у вузький, так і в широкий інтервал. Орієнтири інтерпретації I^2 , наведені у праці Гігінса [Higgins et al., 2003: р. 559]¹, слід також розглядати в контексті оцінювання розміру *пропорції* істинної мінливості, а не її абсолютної величини.

Порівняння показників гетерогенності. Усі описані вище показники гетерогенності базуються на Q (у відношенні до df). Водночас вони призначені для розв'язання різноманітних питань, що робить використання кожного з них виправданим і необхідним. Далі наведено їхню порівняльну характеристику (див. табл. 1).

Довірчі та передбачувальні інтервали

Довірчі інтервали. Оскільки T^2 і I^2 є оцінками відповідних параметрів, а не істинними показниками, то для інтерпретації мінливості величин ефек-

¹ 0% — гетерогенність відсутня, 25% — низька гетерогенність, 50% — середня гетерогенність, 75% — висока гетерогенність.

тів має сенс провести розрахунок довірчих інтервалів, у які із заданою імовірністю мають потрапити істинна мінливість і її пропорція.

Таблиця 1

Показники гетерогенності

Показник і розмах значень	Залежність від розміру вибірки	Залежність від шкали	Розв'язуване питання
$0 \leq Q$	+	-	Чи можна пояснити спостережувану мінливість суто випадковими помилками?
$0 \leq T^2$	-	+	Якою є величина істинної мінливості величин ефектів?
$0 \leq T$	-	+	У який інтервал потрапить більшість величин ефекту?
$0\% \leq I < 100\%$	-	-	Яка пропорція загальної мінливості є істинною, тобто не зумовлена випадковими помилками?

Якщо припустити, що величини ефектів розподілені нормально (відповідає великим вибіркам), стандартну похибку T^2 можна оцінити в такий спосіб.

Спершу обчислити проміжне значення A :

$$2.1. A = \left[df + 2 \left(sw1 - \frac{sw2}{sw1} \right) T^2 + \left(sw2 - 2 \left(\frac{sw3}{sw1} \right) + \frac{(sw2)^2}{(sw1)^2} \right) T^4 \right],$$

де

$$sw1 = \sum_{i=1}^k W_i,$$

$$sw2 = \sum_{i=1}^k W_i^2,$$

$$sw3 = \sum_{i=1}^k W_i^3.$$

Після чого можна знайти V_{T^2} дисперсію T^2 :

$$2.2. V_{T^2} = 2 \times \left(\frac{A}{C^2} \right).$$

Відповідно стандартна похибка (SE_{T^2}) дорівнюватиме $\sqrt{V_{T^2}}$.

Але оскільки розподіл T^2 недостатньою мірою відповідає нормальному, розрахунок конфіденційних інтервалів шляхом множення стандартної похибки на $\pm 1,96$ не приведе до достатньо точного результату доти, доки будуть використовуватися достатньо великі вибірки. Одним із найбільш легких способів розв'язання цієї проблеми є такий:

Якщо $Q > (df + 1)$, знайти B за формулою:

$$2.3. B = 0,5 \times \frac{\ln(Q) - \ln(df)}{\sqrt{2Q} - \sqrt{2 \times df - 1}}$$

Якщо $Q \leq (df + 1)$, скористатися формулою:

$$2.4. B = \sqrt{\frac{1}{2 \times (df - 1) \times \left(1 - \left(\frac{1}{3 \times (df - 1)^2}\right)\right)}}$$

Далі обчислити проміжні значення L та U :

$$2.5. L = \exp\left(0,5 \times \ln\left(\frac{Q}{df}\right) - 1,96 \times B\right)$$

та

$$2.6. U = \exp\left(0,5 \times \ln\left(\frac{Q}{df}\right) + 1,96 \times B\right).$$

Зрештою, довірчі інтервали для істинної дисперсії величин ефектів можуть бути знайдені в такий спосіб:

$$2.7. LL_{T^2} = \frac{df \times (L^2 - 1)}{C}$$

та

$$2.8. UL_{T^2} = \frac{df \times (U^2 - 1)}{C}.$$

При визначенні довірчих інтервалів для стандартного відхилення величин ефектів треба просто знайти квадратний корінь із відповідних значень оцінки дисперсії:

$$2.9. LL_T = \sqrt{LL_{T^2}};$$

$$2.10. UL_T = \sqrt{UL_{T^2}}.$$

Під час побудови довірчих інтервалів для I^2 умови і формули розрахунку B , L і U такі самі, як і у випадку T^2 . Якщо $Q > (df + 1)$, використовують формулу 2.3. Якщо $Q \leq (df + 1)$, використовують формулу 2.4.

Визначення проміжних значень ідентичне формулам 2.5 і 2.6. Далі можна знайти 95-відсотковий довірчий інтервал:

$$2.11. LL_{I^2} = \left(\frac{L^2 - 1}{L^2}\right) \times 100\%$$

та

$$2.12. UL_{I^2} = \left(\frac{U^2 - 1}{U^2}\right) \times 100\%.$$

Будь-яке зі значень для LL і UL , менше за нуль, прирівнюється до нуля. Якщо нижня межа інтервалу перевищує нуль, тоді I^2 має бути статистично

значимим. Утім, оскільки I^2 базується на Q , а вибірковий розподіл Q вивчено краще, ніж вибірковий розподіл I^2 , надійнішим методом оцінювання статистичної значимості I^2 є саме Q -тест.

Передбачувальні інтервали. Дуже часто головною метою метааналізу є обчислення зваженого середнього величин ефектів і його довірчих інтервалів. І хоча це важливе питання, його розв'язання нічого не говорить про розподіл істинних величин ефектів навколо середнього значення. У моделі фіксованих ефектів такий підхід (визначення лише середнього значення і його довірчих інтервалів) є виправданим, оскільки передбачається існування однієї істинної величини ефекту для всіх досліджень. Своєю чергою, в моделі випадкових ефектів необхідно оцінити не тільки істинне середнє величин ефектів, а й їхній розподіл навколо нього. Останнє має стосунок до побудови передбачувальних інтервалів.

Оскільки при побудові довірчих інтервалів необхідно врахувати помилки в оцінюванні величин ефектів і T , розрахунок передбачувальних інтервалів здійснюють у такий спосіб:

$$2.13. LL_{pred} = M^* - t_{df}^{\alpha} \sqrt{T^2 + V_{M^*}},$$

$$2.14. UL_{pred} = M^* + t_{df}^{\alpha} \sqrt{T^2 + V_{M^*}},$$

де M^* — вибіркове зважене середнє значення величин ефектів, T^2 — оцінка істинної дисперсії величин ефектів і V_{M^*} — дисперсія M^* . Фактор t_{df}^{α} — значення з розподілу критичних значень T -Стюдента для відповідної імовірності й кількості ступенів свободи ($df = k - 1$, де k — кількість величин ефектів у вибірці).

Порівняння довірчих і передбачувальних інтервалів зваженого середнього. Важливо пам'ятати, що довірчі та передбачувальні інтервали розв'язують різні питання. Перші призначені для визначення точності оцінювання вибіркового середнього значення величин ефектів, другі — для визначення розкиду всіх можливих величин ефектів навколо вибіркового середнього.

Істотною відмінністю цих двох різновидів інтервалів є те, як вони змінюються мірою збільшення вибірки. Так, довірчий інтервал наближатиметься до нуля, тоді як передбачувальний інтервал буде зменшуватися до певного моменту, після чого зміни практично припиняться. Це пов'язане із тим, що передбачувальний інтервал базується, серед іншого, на T^2 , величина якого не залежить від обсягу вибірки [Borenstein et al., 2009: p. 122–133].

Практичний приклад: стать респондента і його шанси на зайнятість за кордоном

Для прикладу розрахунків показників гетерогенності тут використовуються ті самі емпіричні дані четвертої хвили Європейського соціального дослідження, що й у попередній статті [Дембіцький, 2012: с. 169–172] (див. табл. 2). З міркувань ясності в таблицю включено лише ті показники, що використовуються в поточних розрахунках безпосередньо.

Таблиця 2

Дані, необхідні для аналізу гетерогенності величин ефектів

Країна	W_i	W_i^2	W_i^3	$W_i Y_i$	$W_i Y_i^2$
Бельгія	23,38	546,62	12780,08	10,10	4,36
Болгарія	23,32	543,82	12681,94	4,83	1,00
Швейцарія	26,92	724,69	19508,56	11,09	4,57
Кіпр	13,36	178,49	2384,62	9,26	6,42
Чехія	27,94	780,64	21811,18	7,12	1,82
Німеччина	19,05	362,90	6913,29	14,57	11,15
Данія	16,60	275,56	4574,30	3,02	0,55
Естонія	25,27	638,57	16136,74	31,94	40,37
Іспанія	33,47	1120,24	37494,46	0,00	0,00
Фінляндія	20,04	401,60	8048,10	10,04	5,03
Франція	21,51	462,68	9952,25	13,12	8,00
Велика Британія	27,79	772,28	21461,78	14,90	7,98
Греція	18,64	347,45	6476,46	4,75	1,21
Хорватія	16,51	272,58	4500,30	12,56	9,56
Угорщина	15,17	230,13	3491,06	9,74	6,25
Ірландія	41,94	1758,96	73770,93	12,58	3,77
Ізраїль	27,56	759,55	20933,30	0,28	0,00
Литва	26,12	682,25	17820,48	16,90	10,93
Латвія	25,22	636,05	16041,14	21,54	18,39
Голландія	20,53	421,48	8653,00	16,55	13,34
Норвегія	3,59	12,89	46,27	1,18	0,39
Польща	20,58	423,54	8716,38	14,78	10,61
Португалія	22,33	498,63	11134,38	8,31	3,09
Румунія	19,60	384,16	7529,54	6,02	1,85
Росія	7,62	58,06	442,45	12,73	21,25
Швеція	26,54	704,37	18694,02	13,62	6,98
Словенія	8,80	77,44	681,47	6,27	4,47
Словаччина	29,50	870,25	25672,38	28,85	28,22
Туреччина	7,17	51,41	368,60	-1,69	0,40
Україна	23,00	529,00	12167,00	16,63	12,02
Σ	639,07	15526,32	410886,45	331,58	244,01

Також нагадаю, що середнє зважене значення величин ефектів для відношення шансів (далі — M^*) дорівнює 0,53, а його дисперсія — 0,004. Після здійснення зворотного перетворення (обидва показники розраховано після логарифмічного перетворення) середнє значення дорівнюватиме 1,7, а дисперсія — 1,0.

Передусім необхідно перевірити за допомогою Q -тесту нульову гіпотезу про те, що вся мінливість між вибірковими величинами ефектів спричинена випадковими помилками. Для цього скористаймося формулою 1.2:

$$Q = 244,01 - \frac{331,58^2}{639,07} = 71,97.$$

При кількості ступенів свободи, що дорівнює 29, і рівні α , що дорівнює 0,05, критичне значення розподілу χ^2 дорівнює 42,6. Останнє дає підстави для висновку про гетерогенність величин ефектів.

Здійснюємо оцінювання істинної дисперсії величин ефектів (формула 1.3):

$$df = 30 - 1 = 29,$$

$$C = 639,07 - \frac{15526,32}{639,07} = 614,77,$$

$$T^2 = \frac{71,97 - 29}{614,77} = 0,07.$$

Відповідно стандартне відхилення дорівнюватиме:

$$T = \sqrt{0,07} = 0,26.$$

Тепер розрахуємо довірчий інтервал для T^2 . У випадку припущення про нормальність розподілу величин ефектів спочатку необхідно знайти $sw1$, $sw2$ і $sw3$:

$$sw1 = 639,07$$

$$sw2 = 15526,32$$

$$sw3 = 410886,45$$

Після чого можна розрахувати A (формула 2.1), V_{T^2} (формула 2.2) та SE_{T^2} :

$$A = 29 + \left(639,07 - \frac{15526,32}{639,07} \right) 0,07 + \left(15526,32 - 2 \left(\frac{410886,45}{639,07} \right) + \frac{15526,32^2}{639,07^2} \right) 0,07^2 = 187,40,$$

$$V_{T^2} = 2 \times \left(\frac{187,40}{614,77^2} \right) = 0,00099,$$

$$SE_{T^2} = \sqrt{0,00099} = 0,032.$$

Звідси, 95-відсотковий довірчий інтервал дорівнюватиме:

$$C.i. = \pm 1,96 \times 0,032 = \pm 0,063.$$

Якщо ж нормальність розподілу величин ефектів не передбачається, розрахунки набудуть іншого вигляду. Оскільки $Q = 71,97 > 30 = (df + 1)$, для визначення B скористаймося формулою 2.3:

$$B = 0,5 \times \frac{\ln(71,97) - \ln(29)}{\sqrt{2 \times 71,97} - \sqrt{2 \times 29 - 1}} = 0,10.$$

Проміжні значення (формули 2.5 і 2.6 відповідно):

$$L = \exp \left(0,5 \times \ln \left(\frac{71,97}{29} \right) - 1,96 \times 0,10 \right) = 1,29,$$

$$U = \exp\left(0,5 \times \ln\left(\frac{71,97}{29}\right) + 1,96 \times 0,10\right) = 1,92.$$

Отже, межі 95-відсоткового довірчого інтервалу дорівнюють (згідно із формулами 2.7 і 2.8 відповідно):

$$LL_{T^2} = \frac{29 \times (1,29^2 - 1)}{614,77} = 0,03,$$

$$UL_{T^2} = \frac{29 \times (1,92^2 - 1)}{614,77} = 0,13.$$

Якщо від метрики натурального логарифма перейти до оригінальної метрики відношення шансів, то дисперсія (T^2) буде дорівнювати 1,07, нижній інтервал – 1,03, верхній інтервал – 1,14.

На передостанньому етапі з'ясуємо, яка пропорція загальної мінливості є істинною (формула 1.4):

$$I^2 = \left(\frac{71,97 - 29}{71,97}\right) \times 100\% = 59,7\%.$$

При розрахунку довірчих інтервалів для I^2 значення для L і U будуть такими самими, як і при обчисленні довірчих інтервалів для T^2 . Тому одразу можна перейти до розрахунку меж інтервалу (формули 2.11 і 2.12 відповідно):

$$LL_{I^2} = \left(\frac{1,29^2 - 1}{1,29^2}\right) \times 100\% = 39,9\%,$$

$$UL_{I^2} = \left(\frac{1,92^2 - 1}{1,92^2}\right) \times 100\% = 73,0\%.$$

Зрештою, обчислимо передбачувальні інтервали для величин ефектів (формули 2.13 і 2.14 відповідно), з урахуванням того, що $M^* = 0,53$, $V_{M^*} = 0,004$, а $t_{28}^{0,05} = 1,7$ (кількість ступенів свободи дорівнює 28, рівень $\alpha = 0,05$).

$$LL_{pred} = 0,53 - 1,7 \sqrt{0,070,004} = 0,07,$$

$$UL_{pred} = 0,53 + 1,7 \sqrt{0,070,004} = 0,99.$$

Відповідно, після перетворення на шкалу відношення шансів нижня межа дорівнюватиме 1,07, верхня – 2,69.

Тепер розглянемо всі дані разом (див. табл. 3). Як уже зазначалося раніше, величина $Q(71,97)$ дає підстави дійти висновку про гетерогенність результатів з імовірністю помилки, що не перевищує 5%. Величина істинної мінливості в оригінальній шкалі величин ефектів становить близько 1, що зумовлює майже 60% сукупної мінливості. Остання величина не є достатньо точною і з високою ймовірністю потрапляє в інтервал від 40% до 73%.

Незважаючи на те, що підсумкове середнє значення дорівнює 1,7, що засвідчує слабкий зв'язок між статтю і шансом на зайнятість за кордоном, 95% усіх можливих величин ефектів розподіляються в інтервалі від 1,07 (майже повна відсутність зв'язку) до 2,69 (середній зв'язок).

Підсумкові дані, що описують гетерогенність величин ефектів*

Показник	Значення	Довірчий/передбачувальний інтервал	
		Нижня межа	Верхня межа
M^*	0,53(1,7)	0,41(1,5)	0,65(1,9)
Q	71,97	–	–
T^2	0,07(1,07)	0,03(1,03)	0,13(1,14)
T	0,26(1,3)	0,17(1,19)	0,36(1,43)
I	59,7%	39,9%	73,0%
PI^{**} для M^*	–	0,07(1,07)	0,99(2,69)

* Для випадків, коли показник обчислений у логарифмічній шкалі, в дужках указано значення в одиницях оригінальної метрики величин ефектів (відношення шансів).

** Передбачувальний інтервал.

Висновки

Аналіз гетерогенності є суттєвим елементом коректної інтерпретації результатів метааналізу. Проте у разі крос-національних і крос-культурних досліджень вивчення мінливості взагалі є обов'язковою частиною цього дослідницького підходу, що дає змогу у строгих термінах оцінити й описати відмінності між одиницями аналізу. Однак оцінювання величини істинної мінливості, її пропорції серед сукупної мінливості, побудова передбачувального інтервалу і розрахунок інших показників не є остаточним пунктом аналізу гетерогенності.

Після того, як встановлено гетерогенність результатів і розраховано її показники, розглянуті вище, є два шляхи подальшого аналізу. Перший із них полягає в розподілі величин ефектів на гомогенні підгрупи й аналізі їх, другий – у використанні метарегресії [Leeuw, Нох, 2003: р. 336–339]. Обидва ці інструменти будуть розглянуті в одному з найближчих чисел часопису.

ДОДАТОК

Помилки, яких було припущено у попередній статті

(див.: *Соціологія: теорія, методи, маркетинг.* – 2012. – № 3)

1. На сторінці 162 у сьомому пункті четвертого абзаца замість “синтез ефективних розмірів” має бути “синтез величин ефектів”.

2. На сторінці 168 формула 4.5 повинна мати такий вигляд:

$$C.i. = M \pm 1,96 \times SE_M.$$

3. На сторінці 169 у формулі 5.3 помилково зазначено, що df – це кількість досліджень. Значення df розраховують традиційним способом ($k - 1$), що й показано у формулі 5.5.

4. На сторінці 169 формула 5.4 повинна мати такий вигляд:

$$Q = \sum_{i=1}^k W_i Y_i^2 - \frac{\left(\sum_{i=1}^k W_i Y_i \right)^2}{\sum_{i=1}^k W_i}.$$

Джерела

Дембицкий С. Метаанализ: ключевые понятия и основы вычислений (на примере данных кросс-национальных исследований) / С. Дембицкий // Социология: теория, методы, маркетинг. — 2012. — № 3. — С. 160–174.

Borenstein M. Introduction to Meta-Analysis / M. Borenstein, L. Hedges, J. Higgins, H. Rothstein. — N. J.: Wiley, 2009.

Corcoran J. Meta-Analyses / J. Corcoran, J. Littel // The Handbook of Social Work Research Methods / ed. by B. Thyer. — Los Angeles ; London ; New Delhi ; Singapore ; Washington (DC) : SAGE, 2010. — P. 299–312.

Higgins J. Measuring inconsistency in meta-analyses / J. Higgins, S. Thompson, J. Deeks, D. Altman // British Medical Journal. — 2003. — Vol. 327. — P. 557–560.

Hunter J. Methods of meta-analysis: Correcting error and bias in research findings / J. Hunter, F. Schmidt. — Thousand Oaks ; London ; New Delhi : SAGE, 2004. — 582 p.

Leeuw E. The Use of Meta-Analysis in Cross-National Studies / E. Leeuw, J. Hox // Cross-Cultural Survey Methods / ed. by J. Harkness, F. Van de Vijver, P. Mohler. — N. J.: Wiley, 2003. — P. 329–345.