СОЦИОЛОГИЧЕСКОЕ ОБРАЗОВАНИЕ

УДК 303.447.22

СЕРГЕЙ ДЕМБИЦКИЙ,
кандидат социологических наук, младший научный сотрудник отдела методологии и методов социологии Института социологии НАН Украины

Оценка гетерогенности в метаанализе (на примере данных кросс-национальных исследований)

Аннотация

Статья посвящена основам анализа гетерогенности в рамках метаанализа. Описывается суть оценки гетерогенности, а также рассматриваются следующие вопросы: статистическая проверка с помощью Q-теста; оценка истинной дисперсии величин эффектов с помощью T^2; оценка пропорции наблюдаемой изменчивости, которая указывает на истинную изменчивость величин эффектов с помощью I^2; расчет доверительных и предсказательных интервалов.
На основе практического примера с применением отношения шансов в качестве величин эффектов демонстрируется расчет ключевых показателей гетерогенности (в качестве эмпирической базы использованы результаты четвертой волны Европейского социального исследования).

Ключевые слова: *метаанализ, величина эффекта, модель случайных эффектов, показатели гетерогенности*

Введение

В своей недавней публикации я рассмотрел ключевые понятия и основы вычислений в метаанализе [Дембицкий, 2012]. При этом особое внимание было уделено двум моделям метаанализа — фиксированных и случайных эффектов. В рамках модели случайных эффектов важной задачей является оценка гетерогенности, позволяющая: 1) проанализировать, какая доля изменчивости в величинах эффектов вызвана различиями генеральных совокупностей, из которых они извлечены; 2) сделать прогноз о величине интервала, в который попадает большинство величин эффектов, безотносительно к тому, из какой генеральной совокупности они извлечены; 3) сформулировать гипотезы о факторах изменчивости, то есть тех релевантных факторах, по которым генеральные совокупности отличаются. Последнее открывает возможность раздельного анализа контрастных подгрупп величин эффектов и, соответственно, делает выводы исследования более гибкими за счет привязки к контекстуальным особенностям, важным с точки зрения изучаемой взаимосвязи.

Следует сделать ударение на том, что далее речь пойдет именно об оценке гетерогенности, а не ее проверке. Часто решение о гетерогенности/гомогенности результатов принимается в результате статистической проверки (статистический подход). Однако в рамках данной статьи я отталкиваюсь от позиции Майкла Борнштейна с его соавторами [Borenstein et al., 2009], которые связывают решение о наличии гетерогенности результатов или ее отсутствии с природой анализируемых данных (концептуальный подход). При статистическом подходе принято начинать с использования модели фиксированных эффектов, далее проверять ее результаты на гетерогенность и, в случае ее подтверждения, переходить к модели случайных эффектов. В концептуальном же подходе именно на основании аргументов о природе эмпирического материала (о том, следует ли считать, что величины эффектов извлечены из разных генеральных совокупностей) исследователь выбирает одну из двух моделей. И только в случае выбора модели случайных эффектов проводится оценка гетерогенности, так как она предполагается изначально[1].

Основы оценки гетерогенности

Когда речь идет об оценке гетерогенности в рамках метаанализа, то подразумевается анализ истинной (в противоположность случайным ошибкам) изменчивости величин эффектов. В то же время наблюдаемая изменчивость включает оба компонента — и случайные ошибки, и истинную изменчивость. Механизм, позволяющий разделить эти компоненты, следующий: 1) рассчитать общую величину изменчивости, наблюдаемую между исследованиями, отобранными для анализа; 2) оценить ожидаемую изменчивость для случая, в котором все величины эффектов принадлежат к одной генеральной совокупности (то есть для случая, в котором совокупная изменчивость основывается исключительно на случайных ошибках); 3) найти разницу между общей и ожидаемой изменчивостью (так называемую избыточную вариацию), которая и будет указывать на величину гетерогенности или истинной изменчивости [Borenstein et al., 2009: p. 108].

Статистическая проверка с помощью Q-теста. Традиционно оценка гетерогенности начинается с расчета Q-статистики, демонстрирующей, отличается ли статистически значимо наблюдаемая изменчивость от ожидаемой, основывающейся на случайных ошибках.

---

1 Более подробно о выборе модели метаанализа в контексте оценки гетерогенности результатов см.: [Hunter, Schmidt, 2004: p. 393–399].

Q — показатель, имеющий распределение χ^2, с количеством степеней свободы (df), равным $k - 1$, где k равно количеству величин эффектов [Corcoran, Littel, 2010: p. 303]. Q является стандартизированным показателем, то есть не зависит от метрики используемых величин эффектов. Базовая формула для расчета Q имеет следующий вид:

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

Эквивалентная формула, более пригодная для расчетов:

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

Ожидаемое значение Q для случая, когда все величины эффектов принадлежат одной генеральной совокупности, не отличается от df, поскольку является стандартизированной величиной. Следовательно, истинную вариацию в величинах эффектов можно рассчитать, отняв от наблюдаемого Q его ожидаемое значение, или df. Если эта величина отрицательная или незначительно превышает 0 (достаточна ли величина различия между наблюдаемым и ожидаемым значениями, определяется стандартным для использования критерия χ^2 образом — с помощью p-значения), то проверка не дала статистически значимых результатов. В противном случае принимается решение о гетерогенности, что открывает возможность дальнейшего анализа.

Q-тест является проверкой нулевой гипотезы (согласно которой все исследования описывают один и тот же эффект или принадлежат к одной генеральной совокупности), что связано с проблемами, характерными и для других тестов статистической значимости. Во-первых, если статистически значимые результаты указывают на гетерогенность, то незначимые еще не доказывают ее отсутствия — возможное объяснение может скрываться в малом размере выборки или больших дисперсиях результатов отдельных исследований. Во-вторых, результаты Q-теста используются исключительно для проверки нулевой гипотезы и, соответственно, не являются показателем величины истинной изменчивости. Эту задачу решают другие показатели, о которых будет сказано ниже.

Оценка истинной дисперсии величин эффектов с помощью T^2. Показатель T^2 возвращает оценку изменчивости от стандартизированной шкалы к оригинальной метрике величин эффектов:

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

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

Истинное значение дисперсии величин эффектов не может быть меньше 0, в то время как T^2 может принимать отрицательные значения (при $Q < df$), что связано с ошибками выборки. В таком случае T^2 принимается за 0. Если же $Q > df$, то T^2 будет положительным значением, базирующимся на двух факторах. Первый из них — величина избыточной изменчивости ($Q - df$), второй — размерность шкалы, в которой измерены величины эффектов.

Для оценки стандартного отклонения величин эффектов используется T, которое можно найти простым извлечением квадратного корня из T^2. При нахождении T^2 не требуется предположения о форме распределения величин эффектов, но если есть основания считать ее нормальной, то T можно использовать для нахождения интервала, в который попадет заданная пропорция всех величин эффектов. Например, 95% всех возможных значений попадет в интервал, равный $1{,}96 \times T$.

Вместе с тем такой способ построения распределения величин эффектов оправдан лишь в том случае, когда и эти величины, и T оценены правильно. Поэтому на практике при построении соответствующих интервалов для распределения величин эффектов необходимо принять во внимание ошибки в оценке обоих параметров. Основания соответствующих расчетов рассмотрены ниже.

Оценка пропорции наблюдаемой изменчивости, которая указывает на истинную изменчивость величин эффектов с помощью I^2. I^2 является показателем, производным от Q:

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

Как и Q, показатель I^2 является стандартизированным, то есть не зависит от оригинальной шкалы величин эффектов и изменяется в диапазоне от 0% до 100%. Важной отличительной чертой этой статистики является также то, что она не зависит и от количества анализируемых величин эффектов [Borenstein et al., 2009: p. 109–119].

При этом следует помнить — несмотря на то, что I^2 позволяет оценить пропорцию истинной изменчивости, он не дает никакой другой информации. Так, если соответствующее значение близко к 100%, это еще ничего не говорит о распределении величин эффектов, которые в этом случае могут попадать как в узкий, так и в широкий интервал. Ориентиры интерпретации I^2, приведенные в работе Хиггинса [Higgins et al., 2003: p. 559][1], следует также рассматривать в контексте оценки размера *пропорции* истинной изменчивости, а не ее абсолютной величины.

Сравнение показателей гетерогенности. Все описанные выше показатели гетерогенности базируются на Q (в отношении к df). Вместе с тем они предназначены для решения различных задач, что делает использование каждого из них оправданным и необходимым. Далее приводится их сравнительная характеристика (см. табл. 1).

Доверительные и предсказательные интервалы

Доверительные интервалы. Поскольку T^2 и I^2 являются оценками соответствующих параметров, а не истинными показателями, постольку для интерпретации изменчивости величин эффектов имеет смысл провести рас-

[1] 0% — гетерогенность отсутствует, 25% — низкая гетерогенность, 50% — средняя гетерогенность, 75% — высокая гетерогенность.

чет доверительных интервалов, в которые с заданной вероятностью должны попасть истинная изменчивость и ее пропорция.

Таблица 1

Показатели гетерогенности

Показатель и размах значений	Зависимость от размера выборки	Зависимость от шкалы	Решаемый вопрос
$0 \leq Q$	+	–	Можно ли объяснить наблюдаемую изменчивость исключительно случайными ошибками?
$0 \leq T^2$	–	+	Какова величина истинной изменчивости величин эффектов?
$0 \leq T$	–	+	В какой интервал попадет большинство величин эффекта?
$0\% \leq I < 100\%$	–	–	Какая пропорция общей изменчивости является истинной, то есть не обусловлена случайными ошибками?

Если предположить, что величины эффектов распределены нормально (подходит для больших выборок), стандартную ошибку T^2 можно оценить следующим образом.

Сначала рассчитать промежуточное значение A:

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

где

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

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

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

После чего можно найти V_{T^2} дисперсию T^2:

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

Соответственно стандартная ошибка (SE_{T^2}) будет равна $\sqrt{V_{T^2}}$.

Но поскольку распределение T^2 не достаточно хорошо отвечает нормальному, расчет конфиденциальных интервалов с помощью умножения стандартной ошибки на ±1,96 не приведет к достаточно точному результату до тех пор, пока не будут использоваться достаточно большие выборки. Одним из наиболее легких способов решения этой проблемы является следующий:

Если $Q > (df + 1)$, найти B по формуле:

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

Если $Q \le (df + 1)$, использовать формулу:

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

Далее рассчитать промежуточные значения L и U:

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

и

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

Наконец, доверительные интервалы для истинной дисперсии величин эффектов могут быть найдены следующим образом:

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

и

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

При нахождении доверительных интервалов для стандартного отклонения величин эффектов надо просто извлечь квадратный корень из соответствующих значений оценки дисперсии:

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

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

Во время построения доверительных интервалов для I^2 условия и формулы расчета B, L и U точно такие же, как и в случае T^2. Если $Q > (df + 1)$, используется формула 2.3. Если $Q \le (df + 1)$, используется формула 2.4.

Нахождение промежуточных значений идентично формулам 2.5 и 2.6. Далее можно найти 95-процентный доверительный интервал:

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

и

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

Любое из значений для LL и UL, меньшее нуля, приравнивается к нулю. Если нижняя граница интервала превышает ноль, тогда I^2 должно быть статистически значимым. Однако, поскольку I^2 базируется на Q, а выборочное распределение Q изучено лучше, чем выборочное распределение I^2, более надежным методом оценки статистической значимости I^2 является именно Q-тест.

Предсказательные интервалы. Очень часто главной целью метаанализа является расчет взвешенного среднего величин эффектов и его доверительных интервалов. И хотя это важная задача, ее решение ничего не говорит о распределении истинных величин эффектов вокруг среднего значения. В модели фиксированных эффектов такой подход (нахождение только среднего значения и его доверительных интервалов) является оправданным, поскольку предполагается существование одной и той же истинной величины эффекта для всех исследований. В свою очередь, в модели случайных эффектов необходимо оценить не только истинное среднее величин эффектов, но и их распределение вокруг него. Последнее имеет отношение к построению предсказательных интервалов.

Поскольку при построении доверительных интервалов необходимо учесть ошибки в оценке величин эффектов и T, расчет предсказательных интервалов осуществляется следующим образом:

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

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

где M^* — выборочное взвешенное среднее значение величин эффектов, T^2 — оценка истинной дисперсии величин эффектов и V_{M^*} — дисперсия M^*. Фактор t^{α}_{df} — значение из распределения критических значений t-Стьюдента для соответствующей вероятности и количества степеней свободы ($df = k - 1$, где k — количество величин эффектов в выборке).

Сравнение доверительных и предсказательных интервалов взвешенного среднего. Важно помнить, что доверительные и предсказательные интервалы решают разные задачи. Первые предназначены для определения точности оценки выборочного среднего значения величин эффектов, вторые — для определения разброса всех возможных величин эффектов вокруг выборочного среднего.

Существенным отличием этих двух разновидностей интервалов является то, как они изменяются с увеличением выборки. Так, доверительный интервал будет стремиться к нулю, в то время как предсказательный интервал будет уменьшаться до определенного момента, после чего изменения практически прекратятся. Это связано с тем, что предсказательный интервал базируется, среди прочего, на T^2, величина которого не зависит от размера выборки [Borenstein et al., 2009: p. 122–133].

Практический пример: пол респондента и его шансы на занятость за рубежом

Для примера расчетов показателей гетерогенности используются те же эмпирические данные четвертой волны Европейского социального исследования, что и в предыдущей статье [Дембицкий, 2012: с. 169–172] (см. табл. 2). Из соображений ясности в таблицу включены только те показатели, которые используются в текущих расчетах непосредственно.

Таблица 2

Данные, необходимые для анализа гетерогенности величин эффектов

Страна	W_i	W_i^2	W_i^3	W_iY_i	$W_iY_i^2$
Бельгия	23,38	546,62	12780,08	10,10	4,36
Болгария	23,32	543,82	12681,94	4,83	1,00
Швейцария	26,92	724,69	19508,56	11,09	4,57
Кипр	13,36	178,49	2384,62	9,26	6,42
Чехия	27,94	780,64	21811,18	7,12	1,82
Германия	19,05	362,90	6913,29	14,57	11,15
Дания	16,60	275,56	4574,30	3,02	0,55
Эстония	25,27	638,57	16136,74	31,94	40,37
Испания	33,47	1120,24	37494,46	0,00	0,00
Финляндия	20,04	401,60	8048,10	10,04	5,03
Франция	21,51	462,68	9952,25	13,12	8,00
Великобритания	27,79	772,28	21461,78	14,90	7,98
Греция	18,64	347,45	6476,46	4,75	1,21
Хорватия	16,51	272,58	4500,30	12,56	9,56
Венгрия	15,17	230,13	3491,06	9,74	6,25
Ирландия	41,94	1758,96	73770,93	12,58	3,77
Израиль	27,56	759,55	20933,30	0,28	0,00
Литва	26,12	682,25	17820,48	16,90	10,93
Латвия	25,22	636,05	16041,14	21,54	18,39
Голландия	20,53	421,48	8653,00	16,55	13,34
Норвегия	3,59	12,89	46,27	1,18	0,39
Польша	20,58	423,54	8716,38	14,78	10,61
Португалия	22,33	498,63	11134,38	8,31	3,09
Румыния	19,60	384,16	7529,54	6,02	1,85
Россия	7,62	58,06	442,45	12,73	21,25
Швеция	26,54	704,37	18694,02	13,62	6,98
Словения	8,80	77,44	681,47	6,27	4,47
Словакия	29,50	870,25	25672,38	28,85	28,22
Турция	7,17	51,41	368,60	–1,69	0,40
Украина	23,00	529,00	12167,00	16,63	12,02
Σ	**639,07**	**15526,32**	**410886,45**	**331,58**	**244,01**

Также напомню, что среднее взвешенное значение величин эффектов для отношения шансов (далее — M^*) равно 0,53, а его дисперсия — 0,004. После осуществления обратного преобразования (оба показателя рассчитаны после логарифмического преобразования) среднее значение будет равно 1,7, а дисперсия — 1,0.

Прежде всего необходимо проверить с помощью Q-теста нулевую гипотезу о том, что вся изменчивость между выборочными величинами эффектов вызвана случайными ошибками. Для этого используем формулу 1.2:

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

При количестве степеней свободы, равном 29, и уровне α, равном 0,05, критическое значение распределения χ^2 равно 42,6. Последнее позволяет сделать вывод о гетерогенности величин эффектов.

Произведем оценку истинной дисперсии величин эффектов (формула 1.3):

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

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

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

Соответственно стандартное отклонение будет равно:

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

Теперь рассчитаем доверительный интервал для T^2. В случае предположения о нормальности распределения величин эффектов сначала необходимо найти *sw1, sw2* и *sw3*:

sw1 = 639,07

sw2 = 15526,32

sw3 = 410886,45

После чего можно рассчитать A (формула 2.1), V_{T^2} (формула 2.2) и SE_{T^2}:

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

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

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

Отсюда, 95-процентный доверительный интервал будет равен:

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

Если же нормальность распределения величин эффектов не предполагается, расчеты примут иной вид. Поскольку $Q = 71{,}97 > 30 = (df + 1)$, для нахождения B используем формулу 2.3:

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

Промежуточные значения (формулы 2.5 и 2.6 соответственно):

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

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

Следовательно, границы 95-процентного доверительного интервала равны (согласно формулам 2.7 и 2.8 соответственно):

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

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

Если от метрики натурального логарифма перейти к оригинальной метрике отношения шансов, то дисперсия (T^2) будет равна 1,07, нижний интервал — 1,03, верхний интервал — 1,14.

На предпоследнем этапе выясним, какая пропорция общей изменчивости является истинной (формула 1.4):

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

При расчете доверительных интервалов для I^2 значения для L и U будут точно такими же, как и при расчете доверительных интервалов для T^2. Поэтому сразу же можно перейти к расчету границ интервала (формулы 2.11 и 2.12 соответственно):

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

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

Наконец, вычислим предсказательные интервалы для величин эффектов (формулы 2.13 и 2.14 соответственно), с учетом того, что $M^* = 0{,}53$, $V_{M^*} = 0{,}04$, а $t_{28}^{0,05} = 1{,}7$ (количество степеней свободы равно 28, уровень α — 0,05).

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

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

Соответственно, после преобразования в шкалу отношения шансов нижняя граница будет равна 1,07, верхняя — 2,69.

Теперь рассмотрим все данные вместе (см. табл. 3). Как уже было сказано ранее, величина Q (71,97) позволяет сделать вывод о гетерогенности результатов с вероятностью ошибки, не превышающей 5%. Величина истинной изменчивости в оригинальной шкале величин эффектов составляет около 1, что обусловливает около 60% совокупной изменчивости. Последняя величина не является достаточно точной и с высокой вероятностью попадает в интервал от 40% до 73%.

Несмотря на то, что итоговое среднее значение равно 1,7, что говорит о слабой связи между полом и шансом на занятость за рубежом, 95% всех возможных величин эффектов распределяются в интервале от 1,07 (почти полное отсутствие связи) до 2,69 (средняя связь).

Таблица 3

Итоговые данные, описывающие гетерогенность величин эффектов*

Показатель	Значение	Доверительный/предсказательный интервал	
		Нижняя граница	Верхняя граница
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 случаев, когда показатель измерен в логарифмической шкале, в скобках указаны значения в единицах оригинальной метрики величин эффектов (отношение шансов).

** Предсказательный интервал.

Выводы

Анализ гетерогенности является существенным элементом корректной интерпретации результатов метаанализа. В случае же кросс-национальных и кросс-культурных исследований изучение изменчивости и вовсе является обязательной частью этого исследовательского подхода, позволяющей в строгих терминах оценить и описать отличия между единицами анализа. Однако оценка величины истинной изменчивости, ее пропорции среди совокупной изменчивости, построение предсказательного интервала и расчет других показателей не являются окончательным пунктом анализа гетерогенности.

После того, как установлена гетерогенность результатов и рассчитаны ее показатели, рассмотренные выше, есть два пути дальнейшего анализа. Первый из них заключается в разделении величин эффектов на гомогенные подгруппы и их анализе, второй — в использовании метарегрессии [Leeuw, Hox, 2003: p. 336–339]. Оба этих инструмента будут рассмотрены в одном из ближайших номеров журнала.

ПРИЛОЖЕНИЕ

Ошибки, допущенные в предыдущей статье

(см.: Социология: теория, методы, маркетинг. — 2012. — № 3)

1. На странице 162 в седьмом пункте четвертого абзаца вместо “синтез эффективных размеров” должно быть “синтез величин эффектов”.

2. На странице 168 формула 4.5 должна иметь следующий вид:

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

3. На странице 169 в формуле 5.3 неверно указано, что df — это количество исследований. Значение df рассчитывается традиционным способом ($k - 1$), что и указано в формуле 5.5.

4. На странице 169 формула 5.4 должна иметь следующий вид:

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

Источники

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.