

ПСИХОМЕТРИЧЕСКИЙ АНАЛИЗ ПЕРВОЙ ШКАЛЫ ОПРОСНИКА НЕВРОТИЧЕСКИХ И НЕВРОЗОПОДОБНЫХ РАССТРОЙСТВ НА ОСНОВЕ МОДЕЛИ РАША

Цидик Л. И. (Itsidik@mail.ru)

УО «Гродненский государственный медицинский университет», Гродно, Беларусь

Введение. Учитывая научную несостоятельность классической теории тестов, на основе которой был создан опросник невротических и невротозоподобных расстройств, появилась необходимость применения модели Раша для его модификации. Модель Раша использует процедуры логистической регрессии на основе метода максимального правдоподобия.

Цель. Осуществить анализ конструктивной валидности и надёжности первой шкалы нового многомерного клинического опросника на основе ОНР.

Материал и методы. С помощью методики ОНР были обследованы 286 пациентов, которые находились на стационарном лечении в психоневрологическом отделении.

Результаты. На основе модели Раша была разработана первая шкала модифицированного опросника ОНР, обладающая удовлетворительными психометрическими характеристиками. Трудность пунктов находится в диапазоне от -1,77 до +2,14 логита, утверждения обладают адекватной конструктивной валидностью; индекс надёжности равен 0,91; индекс числа слов равен 5.

Ключевые слова: опросник невротических и невротозоподобных расстройств, модель Раша, конструктивная валидность, надёжность.

Введение

В клинической практике в настоящее время используются психодиагностические методики разной степени стандартизации, которая содержит необходимые теоретические представления и технологические приёмы, позволяющие создавать диагностические методики в клинической психодиагностике и оценивать их эффективность [1].

Большая часть психодиагностических методик, которые используются в клинике, создана и продолжает создаваться на основе классической теории тестов (КТТ). Опросник невротических и невротозоподобных расстройств (ОНР) также был создан на основе данной теории [2]. Классическая теория тестов имеет множество недостатков, основной из которых – зависимость результатов тестирования от нормативной выборки испытуемых [3]. Поэтому появилась необходимость применить для создания новых диагностических инструментов современную теорию тестов, разновидностью которой является модель Раша.

Суть современной теории тестов в том, что вероятность предоставления испытуемым ключевого ответа на задание теста представляет собой функцию от трудности данного задания и выраженности измеряемого свойства и значительное внимание уделяет параметрам не теста в целом, а отдельных диагностических пунктов [4, 5].

Модель Раша использует принципиально иной математический аппарат по сравнению с классической теорией тестов, которая в основном оперирует достаточно простыми в реализации процедурами описательной статистики [6].

Опросник невротических и невротозоподобных расстройств представляет собой адаптированный в Психоневрологическом научно-исследовательском институте имени В. М. Бехтерева

вариант немецкого опросника BVNK-300 [1]. Надёжность, как внутренняя согласованность шкал, не оценивалась. Сведения о конструктивной валидности отсутствуют. Методика не проходила психометрическую проверку на популяции Республики Беларусь.

Целью данного исследования был анализ конструктивной валидности и надёжности первой шкалы нового многомерного клинического опросника, создаваемого на основе опросника невротических и невротозоподобных расстройств. На данном этапе работы произведены расчёт трудности утверждений первой сформированной шкалы опросника, оценка конструктивной валидности методики, надёжность шкалы оценивалась с помощью показателей надёжности и сепарационной статистики на основе модели Раша.

Материал и методы

Исследование проводилось на базе психоневрологического отделения учреждения здравоохранения «Городская клиническая больница № 3 г. Гродно». Осуществлено обследование группы пациентов, страдающих депрессивными расстройствами умеренной степени тяжести (рекуррентное депрессивное расстройство и депрессивный эпизод), невротическими, соматоформными и связанными со стрессом расстройствами (n=286).

Всего обследованы 286 пациентов. Состав группы по полу: мужчины – 129 чел., женщины – 157. Средний возраст испытуемых – 36,4±9,6. Все пациенты, участвовавшие в исследовании, находились на стационарном лечении в психоневрологическом отделении.

Объём выборки для настоящей работы рассчитывался, исходя из требований основного статистического метода работы – модели Раша. Указанная модель позволяет конструировать новые шкалы либо оценивать уже имеющиеся на

выборках относительно небольшого размера.

В процессе обследования испытуемым предлагалось ответить на 300 утверждений опросника: использовался дихотомический вариант заданий. Ответы всех испытуемых на все задания методики были внесены в матрицу данных, которая изначально рассматривалась как единая шкала. Затем производился расчёт конструктивной валидности путём вычисления индексов качества для каждого пункта рассматриваемой шкалы. Пункты, индексы качества которых не входили в диапазон приемлемых значений (для клинических опросников он равен 0,7-1,3), исключались из дальнейших вычислений как нарушающие конструктивную валидность шкалы. Утверждения с низкими значениями по индексам качества располагаются в центре шкалы. Проводилось несколько повторяющихся циклов моделирования шкалы, после каждого из которых оценивались значения индексов качества пунктов и исключались неподходящие. Данные циклы (итерации) повторялись до тех пор, пока значения индексов качества всех оставшихся пунктов не попали в нужный диапазон. Как только это произошло, данная итерация явилась завершающей, а пункты, составляющие шкалу на данном этапе расчётов, стали её окончательным вариантом.

Трудность задания расценивается как доля неключевых ответов от ответов всех испытуемых на конкретное задание. Чем больше значение трудности пункта, тем меньше вероятность предоставления ответа с более высокой оценкой. В модели Раша трудность утверждений шкалы трансформируется в логиты, что позволяет отображать данный параметр на равноинтервальной шкале [7].

Среднее значение шкалы логитов равно 0, в этом случае вероятность ключевого ответа равна вероятности неключевого ответа, а стандартное отклонение равно 1 [8]. Таким образом, пункт с уровнем трудности 0 – это средний по трудности пункт, а пункт с уровнем трудности 1,5 относительно труден. Следует отметить, что для большинства дихотомических заданий трудность варьирует в диапазоне -2 – +2 логита. Задания с трудностью менее -2 логитов оцениваются как чрезмерно лёгкие, задания же с трудностью более +2 логитов – как чрезмерно сложные [9].

Индексы качества характеризуют конструктивную валидность отдельного вопроса, к ним относятся невзвешенное среднее квадратичное – UMS (unweighted mean squared или outfit mean squared), взвешенное среднее квадратичное – WMS (weighted mean squared или infit mean squared). Вычисление указанных индексов основано на анализе стандартизированных остатков. Остатки – математический термин, обозначающий разницу между спрогнозированной моделью вероятностью ключевого ответа на задание и реальным полученным ответом. Стандартизация остатков осуществляется по дисперсии ответов на задание [10].

Модель Раша позволяет оценить надёжность диагностической методики на основе показателя

надёжности и индекса «числа слоёв». Показатель надёжности в модели Раша отображает, насколько наблюдаемая дисперсия данных соответствует истинной дисперсии исследуемого свойства, он концептуально близок коэффициенту альфа-Кронбаха, имеет такие же нормативные диапазоны [2, 3]. Показатель может принимать значения от 0 до 1; значения меньше 0,5 характеризуют надёжность методики как неприемлемую, 0,5-0,6 – плохую, 0,6-0,7 – приемлемую, 0,7-0,9 – хорошую, больше 0,9 – очень хорошую.

Индекс «числа слоёв» представляет собой количество уровней выраженности свойства (статистически значимо отличных друг от друга), которое способен выявить диагностический инструмент в исследуемой выборке, что имеет непосредственное отношение к дифференциально-диагностическим свойствам методики [11].

Результаты и обсуждение

На основе используемого нами статистического метода работы, модели Раша был произведён расчёт трудности пунктов для каждого утверждения шкалы, полученной в результате последней итерации с ошибками измерения, оценена конструктивная валидность утверждений, рассчитаны показатели надёжности и сепарационной статистики.

В таблице 1 представлены рассчитанные с помощью модели Раша трудности части пунктов полученной шкалы. Анализ представленных данных показывает, что трудность утверждений находится в диапазоне от -1,77 до +2,14 логита, а так как наиболее приемлемым для психодиагностических методик является интервал от -2 до +2 логитов, то в данной шкале присутствуют пункты с высоким и низким уровнями трудности, что позволит ей с большой точностью диагностировать разные уровни выраженности исследуемого конструкта.

В метрической системе Раша конструктивная валидность пунктов оценивается на основе значений индексов качества утверждений (UMS и WMS). Индексы качества позволяют оценить, насколько эффективно вопрос функционирует в оценке исследуемого конструкта. Оптимальным значением индексов качества является единица, в таком случае наблюдаемые результаты полностью соответствуют ожидаемым. Нормы индексов качества разнятся в зависимости от конструируемой методики: для клинических опросников диапазон приемлемых значений составляет 0,7-1,3. Значения индексов качества, попадающие в нормативный диапазон, характеризуют соответствующие утверждения как валидные, эффективно оценивающие тот конструкт, который призвана оценить шкала. Утверждения, индексы качества которых выходят за пределы нормативного диапазона, могут быть слишком трудными, плохо сформулированными, или оценивать другой конструкт. Такие утверждения представляют собой угрозу конструктивной валидности шкалы и должны быть элиминированы или переформулированы.

Таблица 1. – Параметры трудности пунктов первой шкалы опросника ОНР (в логитах)

№	Утверждения ОНР	Трудность	Стандартная ошибка
62	Я постоянно чувствую себя усталым	-0,32	0,14
67	В данное время я все делаю без желания	-0,82	0,15
69	Меня угнетает, что я так неуравновешен	-0,28	0,14
74	Мне трудно сосредоточиться и я часто бываю рассеян	-0,46	0,14
77	Даже без большого напряжения я быстро устаю	-0,34	0,14
79	Я нахожу жизнь скучной и мрачной	0,92	0,14
85	Даже после длительного отдыха я не чувствую себя бодрым	-0,21	0,14
87	Без таблеток я засыпаю с большим трудом	0,74	0,14
90	Меня по-настоящему ничто больше не интересует	1,21	0,15
96	Временами я не в состоянии ясно мыслить	-0,05	0,14
101	Я очень быстро устаю	-0,71	0,15
102	Я чувствую разбитость во всем теле, как после большого напряжения	-0,30	0,14
110	У меня часто бывало желание умереть	2,14	0,17
112	Временами я ничего не могу есть, и мне противно думать о еде	0,52	0,14
115	Иногда у меня такое ощущение, как будто я парализован	1,80	0,16
116	Иногда уже с утра я чувствую себя усталым	-1,77	0,17
125	Мне все надоело, мне хотелось бы все бросить	-0,36	0,14
131	Даже когда есть причина, я не могу по-настоящему ничему радоваться	0,20	0,14
136	Я больше уже не способен ни на какие сильные чувства	0,56	0,14
137	Я постоянно чувствую себя «затравленным»	0,94	0,14
143	В данный момент у меня нет никаких желаний, я вял и ко всему безразличен	0,33	0,14
149	Моя умственная работоспособность в последнее время сильно понизилась	-0,59	0,14
151	Уже давно я не могу себя заставить сделать что-либо полезное	0,80	0,14
159	Я уже не в состоянии творчески мыслить	0,60	0,14

Оценка конструктивной валидности пунктов проводилась в несколько повторяющихся циклов моделирования (итераций). Итерации повторялись до тех пор, пока значения индексов качества утверждений не попали в диапазон приемлемых значений. Цикл, в результате которого данное условие было соблюдено, и явился последним. В таблице 2 приведены показатели UMS и WMS, полученные в результате последней и предпоследней итерации. Как видно из таблицы, значение индекса качества п.161 выходит за пределы необходимого диапазона (UMS 1,35) в 55 итерации, что свидетельствует о его несоответствии диагностической направленности шкалы. После его элиминации при проведении заключительной итерации все пункты шкалы вошли в диапазон 0,7-1,3. Такие значения характеризуют утверждения как принадлежащие исследуемому конструкту и эффективно работающие, а шкалу позволяют оценить как достаточную конструктивно валидную.

Таблица 2. – Значения индексов UMS и WMS первой шкалы опросника ОНР

№	Утверждения ОНР	55 итерация		56 итерация	
		WMS	UMS	WMS	UMS
62	Я постоянно чувствую себя усталым	0,98	0,95	0,98	0,94
67	В данное время я все делаю без желания	0,91	0,96	0,90	0,95
69	Меня угнетает, что я так неуравновешен	0,94	0,92	0,95	0,91
74	Мне трудно сосредоточиться, и я часто бываю рассеян	1,05	1,03	1,04	1,01
77	Даже без большого напряжения я быстро устаю	0,83	0,72	0,83	0,72
79	Я нахожу жизнь скучной и мрачной	0,98	0,91	0,98	0,90
85	Даже после длительного отдыха я не чувствую себя бодрым	1,02	1,13	1,02	1,13
87	Без таблеток я засыпаю с большим трудом	1,13	1,11	1,14	1,13
90	Меня по-настоящему ничто больше не интересует	0,92	0,82	0,91	0,81
96	Временами я не в состоянии ясно мыслить	1,03	1,16	1,03	1,13
101	Я очень быстро устаю	1,08	1,12	1,08	1,11
102	Я чувствую разбитость во всем теле, как после большого напряжения	1,15	1,26	1,15	1,27
110	У меня часто бывало желание умереть	0,91	0,76	0,90	0,74
112	Временами я ничего не могу есть, и мне противно думать о еде	1,14	1,22	1,15	1,20
115	Иногда у меня такое ощущение, как будто я парализован	1,02	0,95	1,02	0,94
116	Иногда уже с утра я чувствую себя усталым	1,07	1,00	1,07	1,02
125	Мне все надоело, мне хотелось бы все бросить	0,99	0,91	0,99	0,90
131	Даже когда есть причина, я не могу по-настоящему ничему радоваться	1,07	1,17	1,07	1,15
136	Я больше уже не способен ни на какие сильные чувства	1,02	0,93	1,02	0,93
137	Я постоянно чувствую себя «затравленным»	0,84	0,70	0,84	0,70
143	В данный момент у меня нет никаких желаний, я вял и ко всему безразличен	0,89	0,78	0,88	0,78
149	Моя умственная работоспособность в последнее время сильно понизилась	1,04	1,28	1,04	1,28
151	Уже давно я не могу себя заставить сделать что-либо полезное	1,03	1,17	1,03	1,14
159	Я уже не в состоянии творчески мыслить	0,99	1,14	0,99	1,13
161	Я чувствую себя внутренне беспокойным и взволнованным	1,23	<u>1,35</u>	-	-

Показатели надёжности и сепарационной статистики также рассчитывались отдельно для каждой итерации. Из таблицы 3 видно, что изначально высокое значение индекса надёжности (0,97) при значительном уменьшении количества пунктов в результате проведённых циклов

Таблица 3. – Показатели надёжности и сепарационной статистики первой шкалы опросника ОНР

Статистические показатели	1 итерация	56 итерация
Надёжность	0,97	0,91
Сепарационный индекс	5,4	3,2
Индекс числа слов	7,5	5

моделирования практически не изменилось (0,91). Такое значение показателя характеризует надёжность методики как очень хорошую. Внутренняя согласованность, которую отражает коэффициент надёжности, характеризует устойчивость связей между пунктами диагностической

методики. Индекс числа слоёв значительно изменился и составил 5, что соответствует количеству уровней выраженности свойства, которое способен выявить диагностический инструмент в исследуемой выборке. Данный показатель отражает высокие дифференциально-диагностические свойства шкалы.

Литература

1. Ассанович, М. А. Клиническая психодиагностика : учеб. пособие / М. А. Ассанович. – Минск : Беларусь, 2012. – 343 с.
2. Ассанович, М. А. Статистическое обоснование критериев оценки выраженности измеряемого конструкта в клинической психодиагностике / М. А. Ассанович // Психиатрия, психотерапия и клиническая психология. – 2014. – № 2 (16). – С. 9-18.
3. Фер, Р. М. Психометрика: Введение / Р. М. Фер, В. Р. Бакарарк. – Челябинск : Издат. центр ЮУрГУ, 2010. – 445 с.
4. Crocker, L. Introduction to Classical and Modern Test Theory / L. Crocker, J. Algina. – Mason : Cengage Learning, 2008. – 527 p.
5. McDonald, R. P. Test theory: A unified treatment / R. P. McDonald. – New York ; London : Lawrence Erlbaum Associates, 1999. – 487 p.
6. Coombs, C. H. Psychological scaling without a unit of measurement / C. H. Coombs // Psychol. Rev. – 1950. – Vol. 57, iss. 3. – P. 145-158.
7. Bond, T. G. Applying the Rasch Model. Fundamental Measurement in the Human Sciences / T. G. Bond, C. M. Fox. – New Jersey : Lawrence Erlbaum Associates, 2007. – 348 p.
8. Michell, J. Measurement in psychology. Critical history of a methodological concept / J. Michell. – Cambridge : Cambridge University Press, 1999. – 265 p.
9. Michell, J. An introduction to the logic of psychological measurement / J. Michell. – New York : Lawrence Erlbaum Associates, 1990. – 190 p.
10. Nunnally, J. C. Psychometric theory / J. C. Nunnally, I. H. Bernstein. – 3 rd ed. – New York : McGraw-Hill, 1994. – 736 p.
11. Wright, B. D. Reliability and separation / B. D. Wright // Rasch Measurement Transactions. – 1996. – Vol. 9:4. – P. 472.

References

1. Assanovich MA. Klinicheskaja psihodiagnostika [Clinical psychodiagnosis]. Minsk: Belarus; 2012, 343 p.
2. Assanovich MA. Statistical substantiation of criteria for evaluating the severity of the measured construct in clinical psychodiagnostics. *Psihiatrija, psihoterapija i klinicheskaja psihologija* [Psychiatry, psychotherapy and clinical psychology]. 2014;2(16):9-18 p.
3. Fer RM, Bakarak VR. Psihometrika: Vvedenie [Psychometrics: Introduction]. Cheljabinsk: Izdatelskij centr JuUrGU; 2010, 445 p.
4. Crocker L, Algina, J. Introduction to Classical and Modern Test Theory. Mason: Cengage Learning; 2008. 527 p.
5. McDonald RP. Test theory: A unified treatment. New York, London: Lawrence Erlbaum Associates; 1999. 487 p.
6. Coombs CH. Psychological scaling without a unit of measurement. *Psychol. Rev.* 1950;57:145-158.
7. Bond TG, Fox C M. Applying the Rasch Model. Fundamental Measurement in the Human Sciences. New Jersey: Lawrence Erlbaum Associates; 2007. 348 p.
8. Michell J. Measurement in psychology: Critical history of a methodological concept. Cambridge: Cambridge University Press; 1999. 265 p.
9. Michell J. An introduction to the logic of psychological measurement. Hillsdale. New York: Lawrence Erlbaum Associates; 1990. 190 p.
10. Nunnally JC, Bernstein IH. Psychometric theory. 3 nd ed. New York: McGraw-Hill; 1994. 736 p.
11. Wright BD. Reliability and separation. *Rasch Measurement Transactions*. 1996;9:4. p. 472.

PSYCHOMETRIC ANALYSIS OF THE FIRST SCALE OF THE QUESTIONNAIRE OF NEUROTIC AND NEUROSIS-LIKE DISORDERS BASED ON THE RASCH MODEL

Tsidik L. I.

Educational Institution «Grodno State Medical University», Grodno, Belarus

Background. Given the scientific inconsistency of the classical theory of tests, on the basis of which a questionnaire of neurotic and neurosis-like disorders was developed, the need arose to apply the Rasch model for its modification. The Rasch model uses logistic regression procedures based on the maximum likelihood method.

Aim. To carry out the analysis of the construct validity and reliability of the first scale of a new multidimensional clinical questionnaire based on the BVNK-300.

Material and methods. With the help of the BVNK-300 technique, 286 patients who were under inpatient treatment in the psychoneurological department were examined.

Conclusions. Based on the Rasch model, the first scale of the modified BVNK-300 questionnaire with satisfactory psychometric characteristics has been developed. The difficulty of the items is in the range from -1.77 to +2.14 logits; the statements have adequate construct validity; the reliability index is 0.91; the index of the number of layers is 5.

Keywords: questionnaire of neurotic and neurosis-like disorders, Rasch model, construct validity, reliability.

Поступила: 27.06.2017

Отрецензирована: 30.06.2017