

# Экспериментальное полевое исследование надежности диагнозов МКБ-11 для наиболее обременяющих психических расстройств: результаты исследования на взрослых пациентах в условиях оказания помощи в области психического здоровья 13 стран

Geoffrey M. Reed<sup>1,2</sup>, Pratap Sharan<sup>3</sup>, Tahilia J. Rebello<sup>1,4</sup>, Jared W. Keeley<sup>5</sup>, María Elena Medina-Mora<sup>2</sup>, Oye Gureje<sup>6</sup>, José Luis Ayuso-Mateos<sup>7</sup>, Shigenobu Kanba<sup>8</sup>, Brigitte Khoury<sup>9</sup>, Cary S. Kogan<sup>10</sup>, Valery N. Krasnov<sup>11</sup>, Mario Maj<sup>12</sup>, Jair de Jesus Mari<sup>13</sup>, Dan J. Stein<sup>14</sup>, Min Zhao<sup>15</sup>, Tsuyoshi Akiyama<sup>16</sup>, Howard F. Andrews<sup>4,17</sup>, Elson Asevedo<sup>13</sup>, Majda Cheour<sup>18</sup>, Tecelli Domínguez-Martínez<sup>2,19</sup>, Joseph El-Khoury<sup>9</sup>, Andrea Fiorillo<sup>12</sup>, Jean Grenier<sup>20</sup>, Nitin Gupta<sup>21</sup>, Lola Kola<sup>6</sup>, Maya Kulygina<sup>11</sup>, Itziar Leal-Leturia<sup>7</sup>, Mario Luciano<sup>12</sup>, Bulumko Lusu<sup>14</sup>, J. Nicolas I. Martínez-López<sup>2</sup>, Chihiro Matsumoto<sup>22</sup>, Lucky Umukoro Onofa<sup>23</sup>, Sabrina Paterniti<sup>24</sup>, Shivani Purnima<sup>3</sup>, Rebeca Robles<sup>2</sup>, Manoj K. Sahu<sup>25</sup>, Goodman Sibeko<sup>14</sup>, Na Zhong<sup>15</sup>, Michael B. First<sup>1,4</sup>, Wolfgang Gaebel<sup>26</sup>, Anne M. Lovell<sup>27</sup>, Toshimasa Maruta<sup>28</sup>, Michael C. Roberts<sup>29</sup>, Kathleen M. Pike<sup>1</sup>

<sup>1</sup>Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY, USA, <sup>2</sup>National Institute of Psychiatry Ramón de la Fuente Muñiz, Mexico City, Mexico, <sup>3</sup>Department of Psychiatry, All India Institute of Medical Sciences, New Delhi, India, <sup>4</sup>New York State Psychiatric Institute, New York, NY, USA, <sup>5</sup>Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA, <sup>6</sup>Department of Psychiatry, University of Ibadan, Nigeria, <sup>7</sup>Department of Psychiatry, Universidad Autonoma de Madrid, IIS-P and Instituto de Salud Carlos III, Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, Spain, <sup>8</sup>Department of Neuropsychiatry, Kyushu University, Fukuoka City, Japan, <sup>9</sup>Department of Psychiatry, American University of Beirut Medical Center, Beirut, Lebanon, <sup>10</sup>School of Psychology, University of Ottawa, Ottawa, Ontario, Canada, <sup>11</sup>Moscow Research Institute of Psychiatry, National Medical Research Centre for Psychiatry and Narcology, Moscow, Russian Federation, <sup>12</sup>Department of Psychiatry, University of Campania "L. Vanvitelli", Naples, Italy, <sup>13</sup>Department of Psychiatry, Universidade Federal de São Paulo, São Paulo, Brazil, <sup>14</sup>Department of Psychiatry, University of Cape Town and South African Medical Research Council Unit on Risk and Resilience in Mental Disorders, Cape Town, South Africa, <sup>15</sup>Shanghai Mental Health Center and Department of Psychiatry, Shanghai Jiao Tong University School of Medicine, Shanghai, People's Republic of China, <sup>16</sup>NTT Medical Center Tokyo, Tokyo, Japan, <sup>17</sup>Departments of Biostatistics and Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY, USA, <sup>18</sup>Department of Psychiatry, Tunis Al Manar University and Al Razi Hospital, Tunis, Tunisia, <sup>19</sup>Catedras CONACYT, National Council for Science and Technology, Mexico City, Mexico, <sup>20</sup>Institut du Savoir Montfort – Hôpital Montfort & Université d'Ottawa, Ottawa, Ontario, Canada, <sup>21</sup>Department of Psychiatry, Government Medical College and Hospital, Chandigarh, India, <sup>22</sup>Japanese Society of Psychiatry and Neurology, Tokyo, Japan, <sup>23</sup>Federal Neuropsychiatric Hospital Aro, Abeokuta, Nigeria, <sup>24</sup>Institute of Mental Health Research, Royal Ottawa Mental Health Centre, and Department of Psychiatry, University of Ottawa, Ontario, Canada, <sup>25</sup>Pt. Jawahar Lal Nehru Memorial Medical College, Raipur, Chhattisgarh, India, <sup>26</sup>Department of Psychiatry and Psychotherapy, Medical Faculty, Heinrich-Heine University, Düsseldorf, Germany, <sup>27</sup>Institut National de la Santé et de la Recherche Médicale U988, Paris, France, <sup>28</sup>Health Management Center, Seitoku University, Matsudo City, Japan, <sup>29</sup>Office of Graduate Studies and Clinical Child Psychology Program, University of Kansas, Lawrence, KS, USA

Перевод: Красавин Г.А. (Москва)

Редактура: к.м.н. Федотов И.А. (Рязань)

*Надежная, клинически полезная и повсеместно применимая классификация психических расстройств – необходимый фундамент психического здоровья во всем мире. Всемирная организация здравоохранения заканчивает работу над Международной классификацией болезней и проблем, связанных со здоровьем, 11-го пересмотра (МКБ-11). Настоящее исследование оценивает межэкспертную надежность диагностики психических расстройств, ответственных за наибольшую долю глобального бремени и самый высокий уровень потребления услуг: шизофрении и первичных психотических расстройств, расстройств настроения, тревожных и связанных со страхом расстройств, а также расстройств, связанных со стрессом, среди взрослых пациентов, проходивших лечение в 28 центрах-участниках из 13 стран мира. Был использован дизайн исследования с действующими совместно объединенными экспертами с особенной фокусировкой на том, соглашались ли два клинициста, полагаясь на одинаковую клиническую информацию, с определенным диагнозом, когда независимо применяли диагностическое руководство к МКБ-11. Всего 1806 пациентов были оценены 339 клиницистами, используя региональные языки. Внутрикласовый коэффициент каппа для диагнозов, взвешенный по преобладанию местоположения и исследования, находился в интервале от 0,45 (дистимическое расстройство) до 0,88 (социофобия) и считался от умеренного до почти идеального для всех диагнозов. В целом, надежность диагностических руководств к МКБ-11 была выше, чем описываемая ранее для эквивалентных руководств к МКБ-10. Эти данные поддерживают приемлемость внедрения диагностических руководств к МКБ-11 на глобальном уровне. Эти выводы будут сообщаться перед публикацией следующих пересмотров диагностических руководств к МКБ-11 и программ развития, чтобы помочь обучению специалистов и внедрению МКБ-11 в страны – члены ВОЗ.*

**Ключевые слова:** Международная классификация болезней, МКБ-11, диагностика, психические расстройства, надежность, шизофрения, аффективные расстройства, тревожные расстройства, связанные со стрессом расстройства.

(World Psychiatry. 2018;17(2):174-186)

Классификация, которая обеспечивает удовлетворительную надежность, клиническую пригодность и повсеместную применимость диагнозов психических расстройств, является необходимым фундаментом психического здоровья во всем мире. Такая система способствует эффективной идентификации людей с наибольшими потребностями в области сохранения психического здоровья в то время, когда они ищут медицинскую помощь, а также обеспечивает доступ к надлежащей и экономически эффективной помощи<sup>1</sup>.

Классификации формируют связь между потребителями медицинской помощи и информацией о здоровье, а также являются немаловажным основанием для принятия решений в области стратегий здравоохранения и распределения ресурсов на системном, национальном и глобальном уровнях. Классификация, которой слишком сложно пользоваться или которая не дает клинически значимую информацию для занимающихся лечением профессионалов в области здравоохранения, не будет использоваться как задумано, она не может предоставлять достоверные статистические данные и не будет способствовать надлежащей клинической практике, исследованию и выработке тактики<sup>2</sup>.

Всемирная организация здравоохранения (ВОЗ) завершает создание Международной классификации болезней и проблем, связанных со здоровьем, 11-го пересмотра (МКБ-11), которая должна быть обнародована с целью применения в странах – членах ВОЗ в 2018 г. Данный пересмотр дал хорошую возможность привести классификацию психических расстройств и расстройств поведения по МКБ в соответствие с текущими эмпирическими данными и клиническим опытом.

Для достижения этих целей Департамент по вопросам психического здоровья и зависимостям ВОЗ утвердил ряд рабочих групп, чтобы сфокусироваться на расстройствах из разных областей, и эти группы провели всестороннюю экспертизу доступных данных, на которые, собственно, и опираются их рекомендации<sup>3–8</sup>. Вдобавок Департамент по вопросам психического здоровья и зависимостям выполнил систематическую и всеобъемлющую программу, формирующую и оценивающую полевые исследования, направленные, в частности, на повсеместную применимость и клиническую полезность пересмотренных клинических описаний и диагностических указаний (КОДУ), касающихся психических, поведенческих расстройств и расстройств психологического развития, для МКБ-11. Эти полевые исследования существенно отличаются от предыдущих попыток тем, что в них используются инновационные методики с целью изучения применения этих руководств в контексте клинического принятия решений<sup>9</sup>.

В настоящий момент в ряде базирующихся на интернет-технологиях многоязычных полевых исследований типа случай-контроль сравнивается эффективность КОДУ к МКБ-11 относительно КОДУ к МКБ-10<sup>10</sup> в образовании более последовательных клинических суждений с использованием стандартизированных материалов в форме коротких описаний, поскольку это позволяет эмпирически управлять конкретной клинической информацией так, чтобы обособить влияние этой классификационной системы на диагностическое принятие решений<sup>11,12</sup>. Использование коротких письменных описаний дает много преимуществ в плане стандартизации и контроля за исследованием.

Следующим важным шагом в оценке КОДУ являются исследования их внедрения в качестве клинической позиции, что обеспечивает более полное приближение искусности в оценке, интерпретации и принятии решений, связанных с постановкой диагноза реальным пациентам. Поэтому, чтобы изучить диагностическую надежность и клиническую полезность предлагаемых к МКБ-11 КОДУ, в ряде стран в клинических условиях проводятся полевые иссле-

дования внедрения в среду (ПИВС). Центры ПИВС находятся в странах, которые сильно различаются языком, культурой и уровнем ресурсов. Начальные результаты этих исследований описаны в настоящей работе.

Предлагаемая структура и содержание КОДУ к МКБ-11 были избраны, чтобы улучшить их клиническую полезность, валидность и надежность<sup>13</sup>. ВОЗ сосредоточилась на улучшении клинической применимости МКБ-11, потому что это очень важно в соответствии с целями ВОЗ в сфере здравоохранения, связанными со снижением глобального бремени психических расстройств<sup>1</sup>. ВОЗ определяет клиническую полезность для классификационных систем как то, что относится к их значению в коммуникации между заинтересованными сторонами, к их особенностям внедрения в клиническую практику (например, хорошее качество формы, время, необходимое для их применения) и к их полезности в плане принятия организационных клинических решений<sup>14</sup>.

Таким образом, клиническая полезность, валидность и надежность — разные, но перемежающиеся конструкции<sup>15</sup>. В ходе полевых испытаний КОДУ к МКБ-10 был пример взаимоотношения между надежностью и клинической полезностью диагнозов, который показал, что диагнозы с меньшей надежностью имели меньшую, чем в среднем, оценку клинической применимости (например, соответствие диагноза, уверенность в диагнозе, легкость применения и адекватность описания). Схожим образом аспекты валидности диагностических конструкций связаны со свойственной им клинической применимостью в уходе за пациентами, например в предсказании терапевтического ответа или прогноза заболевания<sup>17</sup>.

Надежность диагнозов психических расстройств была в центре внимания в ходе процесса пересмотра как МКБ, так и диагностического и статистического руководства по психическим расстройствам (DSM) американской психиатрической ассоциации с 70-х годов прошлого века<sup>18</sup>. Обе классификационные системы адаптировали описательный подход в предоставлении диагностических указаний<sup>19–21</sup>, частично основываясь на исследованиях, говорящих о том, что недостатки в классификационных системах до DSM-III были одним из основных источников ненадежности<sup>22–24</sup>.

В общем, исследования надежности диагностических классификаций после публикации DSM-III задокументировали улучшение результатов<sup>18</sup>. Тем не менее самая низкая диагностическая надежность была задокументирована в полевых исследованиях DSM-5<sup>25</sup>, что, по сравнению с предыдущими полевыми исследованиями, подчеркивает чрезвычайное влияние методологии на предполагаемую диагностическую надежность<sup>26</sup>. Таким образом, надежность – не только свойство конкретной классификации, но еще и продукт методов, используемых для того, чтобы оценить ее. Это делает довольно сложным сравнение исследований с разной методологией.

Данное исследование имеет натуралистический дизайн со связанными экспертами, чтобы оценить межэкспертную надежность. В отличие от предыдущих исследований надежности классификационных систем<sup>27,28</sup>, структурированные интервью, с помощью которых предполагалось значительно повысить надежность<sup>29–31</sup>, не использовались. Клинические эксперты не получали никаких инструкций и никакого обучения относительно того, как им следует проводить диагностическое интервью, и они прошли лишь минимальное обучение руководства к новой МКБ-11. Эта попытка была предпринята, чтобы приблизиться к условиям, при которых данное руководство будет применяться в клинической практике после его публикации.

Чтобы минимизировать расхождения и особо сконцентрироваться на вопросе, примут ли два клинициста, основываясь на одинаковой клинической информации, одинаковые

диагностические решения по конкретному пациенту, независимо применяя диагностическое руководство к МКБ-11, был избран дизайн исследования с использованием связанных экспертов.

Так же как и в натуралистическом дизайне текущих ПИВС МКБ-11, разработчики полевых исследований для DSM-III, МКБ-10 и DSM-5 в свое время решили не использовать структурированные диагностические интервью, потому что они нечасто используются в обычных клинических условиях<sup>16,25,32,33</sup>. Полевые исследования DSM-III и КОДУ к МКБ-10 показали хорошую диагностическую надежность большинства классов расстройств. Тем не менее предполагаемая надежность была, вероятно, завышена в случае с DSM-III за счет представления прогнозов только для групп расстройств (а не индивидуальных диагнозов)<sup>32</sup> и за счет проведения клинических разборов, на которых один диагност опрашивал конкретного пациента, а затем докладывал этот случай другим экспертами, чтобы создать межэкспертную надежность в случае с КОДУ к МКБ-10<sup>16</sup>.

В полевых исследованиях DSM-5 также использовался натуралистический дизайн, задействующий двух диагностов, которые должны были оценивать согласие между экспертами и высчитывать надежность на уровне индивидуальных диагнозов<sup>34</sup>. Однако те полевые исследования использовали последовательный тест-ретест дизайн (два диагноста опрашивали участника в разное время), чтобы установить межэкспертную надежность, а не одновременный дизайн с соединенными экспертами (два клинициста опрашивают пациента вместе), использованный в ПИВС МКБ-11. Дизайн DSM-5 по этой причине не имел регулирования отклонений информации и поэтому, почти без сомнений, давал бы более низкую надежность<sup>26,35</sup>. Таким образом, прогнозируемая надежность текущих полевых исследований DSM-5 и текущих ПИВС МКБ-11 несравнима. Спорно, что дизайн DSM-5 в большей степени является проверкой диагностической надежности психиатрических диагнозов, а не конкретно нового диагностического руководства.

ПИВС МКБ-11 были созданы как экспериментальные исследования с целью использования результатов в финальном пересмотре руководств, а не как просто оценочные полевые исследования, которые были нацелены на определение того, что пользователи могут ожидать исходя из психометрических свойств классификации, после того как работа над классификацией будет закончена<sup>36</sup>. Одновременный дизайн надежности со связанными экспертами был более предпочтительным в ПИВС, так как он позволял сконцентрироваться на разнообразии в применении или интерпретации диагностических руководств, регулируя отклонения благодаря факторам, связанным с пациентами (например, рассказывающими диагностам разные истории) или внешним факторам, связанным с клиницистом (таким как различия в тщательности опроса).

Одновременный дизайн со связанными экспертами применялся в ПИВС МКБ-11. Особое внимание уделялось роли диагностических руководств самих по себе как источника ненадежности. В экспериментальных полевых исследованиях определение высоких уровней несоместимости критериев, связанных с клиницистами, должно было побуждать изменения в диагностических руководствах, в то время как ошибки клиницистов с большей вероятностью были бы связаны с обучением использованию классификации и клинического опроса.

Изучение надежности ПИВС, описанное в этой работе, в первую очередь было нацелено на четыре группы расстройств у взрослых пациентов: шизофрения и другие первичные психотические расстройства, аффективные расстройства (включая как депрессивное, так и биполярное

расстройства), тревога и связанные со страхом расстройства, а также расстройства, связанные со стрессом. Эти диагнозы отвечают за большую долю глобального бремени болезни в ряду психических расстройств<sup>37</sup>, а также за наиболее высокие уровни использования услуг в условиях здравоохранения в области психического здоровья.

В настоящей работе описаны результаты ПИВС, касающиеся надежности предлагаемых для МКБ-11 КОДУ и полученные на взрослых пациентах в 13 странах.

## МЕТОДЫ

### Дизайн исследования и порядок действий

Для того чтобы исследовать надежность предлагаемых диагностических руководств к МКБ-11, были выполнены два протокола исследования. Протокол 1 проверял надежность руководств для шизофрении и других первичных психотических расстройств и аффективных расстройств, в то время как Протокол 2 – руководства для аффективных расстройств, тревожных и связанных со страхом расстройств, а также расстройств, связанных со стрессом.

Для Протокола 1 подходили взрослые ( $\geq 18$  лет) пациенты, испытывающие любые психотические симптомы и получающие лечение в центре, который участвует в полевом исследовании. Для Протокола 2 подходили взрослые ( $\geq 18$  лет) пациенты, испытывающие аффективные, тревожные и связанные со стрессом (но не психотические) симптомы и находящиеся на лечении в центрах, которые участвуют в полевом исследовании. Эти требования предназначались для того, чтобы сделать расширенную выборку, в которой с большой долей вероятности было хоть одно состояние, подлежащее тестированию, но диагностический статус которого не был бы определен заранее.

Критериями исключения для обоих протоколов были: сложности во взаимодействии, достаточные для того, чтобы мешать пациенту участвовать в диагностическом интервью (например, недостаточное умение говорить на языке, на котором говорят клиницисты); когнитивные нарушения в такой мере, в которой это мешает участию в диагностическом интервью; текущее нарушение дееспособности из-за тяжелого соматического заболевания или боли, текущая интоксикация психоактивными веществами, или синдром отмены, или существенные побочные эффекты лекарств; текущая высокая опасность причинения вреда себе или другим. Эти критерии, по существу, были нужны для того, чтобы допустить любых согласных пациентов, имеющих указанные симптомы, к участию в исследовании, исключая только тех, кто не может участвовать в диагностическом интервью по веским причинам.

Протоколы были внедрены в 28 центрах в 13 странах. Дополнительная информация о центрах представлена в табл. 1.

Для диагностических интервью всегда использовался местный язык. Руководство к МКБ-11, образовательные материалы и все материалы для исследования были разработаны на английском языке. Затем материалы были переведены на другие языки: китайский, японский, русский и испанский – при участии полевых исследовательских центров, используя обратный перевод. В других центрах руководства и обучающие материалы на английском языке использовались, даже если интервью проводилось на другом языке, снова моделируя обстоятельства при которых будет внедряться МКБ-11.

Все центры перед началом исследования получали одобрение от этических комитетов местных учреждений. Исследовательские команды устанавливали местные процедуры для получения согласия и для отчетов и обращений по поводу неблагоприятных событий, которые могли возникнуть у участников, когда они опрашивались бы в рамках исследования (например, невозможность завершения

Таблица 1. Участвующая страна и информация об исследовательском центре				
Страна	Выполняемый(ые) протокол(ы)	Количество центров	Названия центров	Количество экспертов
Бразилия	1	1	Universidade Federal de São Paulo	21
Канада	2	1	Royal Ottawa Mental Health Centre/University of Ottawa Institute of Mental Health Research	7
Китай	1 и 2	1	Shanghai Mental Health Center	25
Индия	1 и 2	3	All India Institute of Medical Sciences	44
			New Delhi Government Medical College Hospital	
			Chandigarh Pandit Jawaharlal Nehru Memorial Medical College, Raipur	
Италия	1	1	University of Campania "L. Vanvitelli", Naples	14
Япония	1 и 2	11	Kyushu University	90
			Hokkaido University	
			University of Occupational & Environmental Health, Kitakyushu	
			Tokyo Medical Dental University	
			Tokyo Metropolitan Matsuzawa Hospital	
			Nihon University School of Medicine, Tokyo	
			Nagoya University	
			Hizen National Psychiatric Center, Yoshinogari	
			NTT Medical Center Tokyo	
			Tokyo University	
Tokushima University				
Ливан	1 и 2	2	American University of Beirut	14
			Hôpital Psychiatrique De La Croix, Jal El Dib	
Мексика	1 и 2	1	National Institute of Psychiatry Ramón de la Fuente Muñiz, Mexico City	23
Нигерия	1	2	University College Hospital, Ibadan	32
			Federal Neuropsychiatric Hospital, Aro, Abeokuta	
Россия	1	2	Moscow Research Institute of Psychiatry	41
			First Saint Petersburg City Mental Hospital	
Южная Африка	1 и 2	1	Valkenberg Psychiatric Hospital, Cape Town	10
Испания	1 и 2	1	Hospital Universitario La Princesa, Madrid	6
Тунис	1 и 2	1	Razi Hospital, Tunis	12

интервью из-за сильной тяжести симптомов или дистресса). Участникам были присвоены уникальные идентификационные номера, и никакая конфиденциальная или идентифицирующая информация не сообщалась никому за пределами центра.

За набор клинических экспертов в каждом центре отвечал его директор. Все клинические эксперты были квалифицированы согласно медицинским стандартам своих стран и имели право выставлять диагнозы психических расстройств независимо, в качестве части своей сферы деятельности. Успешно обучающиеся по специальности «психиатрия» ординаторы (после окончания первых двух лет ординатуры) могли работать как интервьюеры, но они всегда работали в паре с полностью квалифицированным специалистом. Обучение было организовано или на уровне центра, или на уровне нескольких центров в некоторых странах.

Клинические эксперты были обеспечены тестируемым диагностическим руководством к МКБ-11, и их попросили просмотреть руководство до прохождения обучения. Во время учебных занятий рассматривались главные особенности диагностических руководств к МКБ-11 в тех областях, которые входили в протоколы, и их отличия от МКБ-10. На занятиях использовался стандартный набор слайдов, разработанных ВОЗ. Интерактивные упражнения давали

возможность попрактиковаться в применении руководства к кратким описаниям случаев. Единственная разница между Протоколами 1 и 2 состояла в том, что для первого клинические эксперты были проинформированы о том, что им будет необходимо оценивать шизофрению и другие первичные психотические и аффективные расстройства, наряду с любыми другими областями, которые они считают подходящими для формулирования диагноза, хотя для последнего им было необходимо исследовать аффективные, тревожные и связанные со страхом расстройства, а также расстройства, связанные со стрессом. Других инструкций по поводу проведения интервью не давалось, и выбор того, как лучше осуществлять экспертизу, был отдан клиническим экспертам, согласно их профессиональной подготовке и ежедневной практике, в том виде, в каком это будет внедряться в МКБ-11.

Обучающие занятия длились примерно два часа для каждого протокола (т. е. примерно четыре часа для центров, где проводились оба протокола – 1 и 2). Обучающие занятия также не отличались от тех, которые клиницисты на самом деле получают, когда МКБ-11 будет внедряться в их странах. Темы этих занятий также включают в себя ход исследования и процедуры сбора материала. После обучения и до того, как начать собирать данные, клинические эксперты реги-

Показатель	Всего (n=339)	Бразилия (n=21)	Канада (n=7)	Китай (n=25)	Индия (n=44)	Италия (n=14)	Япония (n=90)	Ливан (n=14)	Мексика (n=23)	Нигерия (n=32)	Россия (n=41)	Южная Африка (n=10)	Испания (n=6)	Тунис (n=12)
Возраст, годы (среднее±СО)	37,2±8,3	35,5±8,6	44,4±13,8	32,6±4,8	34,1±7,4	39,8±6,2	38,9±7,7	36,1±8,6	37,6±7,9	37,8±6,0	39,2±11,1	35,5±7,0	32,0±5,9	38,3±9,3
Пол, n (%):														
мужчины	192 (56,6)	10 (47,6)	1 (14,3)	5 (20,0)	29 (65,9)	9 (64,3)	72 (80,0)	7 (50,0)	12 (52,2)	25 (78,1)	14 (34,1)	5 (50,0)	2 (33,3)	1 (8,3)
женщины	145 (42,8)	11 (52,4)	6 (85,7)	20 (80,0)	15 (34,1)	5 (35,7)	16 (17,8)	7 (50,0)	11 (47,8)	7 (21,9)	27 (65,9)	5 (50,0)	4 (66,7)	11 (91,7)
Клиническая специальность, n (%):														
психиатрия	316 (93,2)	21 (100)	2 (28,6)	25 (100)	44 (100)	14 (100)	88 (97,8)	11 (78,6)	22 (95,7)	32 (100)	39 (95,1)	3 (30,0)	5 (83,3)	10 (83,3)
психология	13 (3,8)	0	5 (71,4)	0	0	0	0	3 (21,4)	1 (4,3)	0	1 (2,4)	1 (10,0)	1 (16,7)	1 (8,3)
сестринское дело	5 (1,5)	0	0	0	0	0	0	0	0	0	0	5 (50,0)	0	0
другая специальность	5 (1,5)	0	0	0	0	0	2 (2,2)	0	0	0	1 (2,4)	1 (10,0)	0	1 (8,3)
Опыт работы, годы (среднее±СО)	7,6±7,5	6,6±7,4	13,3±11,9	4,2±3,9	5,4±6,4	7,7±7,2	7,1±6,7	7,8±7,4	9,2±8,2	5,8±4,1	13,6±10,3	6,4±3,9	6,0±4,8	6,5±6,6

стрировались, чтобы участвовать, используя систему онлайн-регистрации, предоставляя демографическую информацию наряду с деталями, касающимися их клинического опыта (табл. 2).

Более широкую группу клиницистов в каждом исследовательском центре проинформировали о критериях включения и исключения для исследования, а также о процедурах по направлению к врачу и попросили у них направлять пациентов или на Протокол 1, или на Протокол 2. В большинстве центров клинические эксперты, проводившие совместные интервью, также были частью пула направляющих клиницистов, но в этом случае им нельзя было опрашивать пациента, которого они направили. Направляющие клиницисты были приглашены к участию в обучающих занятиях для интервьюеров, хотя это не было обязательным.

После направления координатор исследования объяснял суть исследования направленным пациентам и брал от них информированное согласие. После получения информированного согласия пациентов опрашивали два клинициста, которые до этого не контактировали с пациентом в клиническом плане. Один клинический эксперт являлся ведущим интервьюером, а другой – наблюдателем. Наблюдатель мог задавать дополнительные вопросы в конце интервью. Пары клиницистов менялись, насколько это было возможно, учитывая разные ограничения в доступности и в составлении графика работы, а также участвующие клиницисты менялись ролями ведущего интервьюера и наблюдателя.

Клинические эксперты были проинструктированы на счет того, чтобы выделять 60–90 минут на одно совместное интервью. Их также попросили проводить обследование так, как они бы это делали в своей обычной практике. Детализированность и длительность диагностических интервью были значительно согласованы с обычной практикой в участвующих центрах психического здоровья.

Основываясь на интервью и, в некоторых случаях, на дополнительных вспомогательных материалах, предоставляемых клиницистами (например, медицинской карте пациента, исключая текущие или более ранние психиатрические диагнозы и назначаемые психотропные препараты,

опросе членов семьи), клинические эксперты независимо приходили к формулировке диагноза, который мог включать до трех отдельных диагнозов. Диагнозы не имели иерархии (т. е. не было определенного основного, вторичного или третичного диагноза) и могли выпадать на любую диагностическую группу: психическую, поведенческую или группу нарушений психического развития. Участвующие клиницисты также могли указывать диагнозы не психических или поведенческих расстройств или вообще не указывать диагноз. Для диагнозов, входящих в Протоколы 1 и 2, задавались подробные дополнительные вопросы о наличии симптомов и клинической применимости руководства.

Следом за интервью оба клинических эксперта независимо предоставляли данные, основанные на интервью, используя безопасную сетевую систему сбора данных. Участвующие клиницисты были проинструктированы о том, что они должны записывать данные в течение 24 часов. Предоставляемая информация включала в себя сформулированный каждым клиническим экспертом диагноз, оценку наличия или отсутствия каждого элемента любого расстройства из диагностических групп, которое было в центре внимания Протокола 1 или Протокола 2. Предоставленные каждому клиницистом данные также включали ответы на подробные вопросы о клинической применимости диагностических руководств, сформулированные по результатам применения к конкретному пациенту.

### Участники

Всего в качестве клинических экспертов по Протоколу 1 и/или Протоколу 2 работали 339 клиницистов из 28 исследовательских центров в 13 странах (см. табл. 2). Средний возраст клинических экспертов составлял 37,2±8,3 года, средний возраст по странам был сопоставим. Мужской пол несколько преобладал среди клинических экспертов в общей выборке (56,6%). Подавляющее большинство среди клинических экспертов составляли психиатры (93,2%), с небольшим представительством психологов (3,8%), медицинских сестер (1,5%) и других медицинских специалистов (1,5%). Средний клинический опыт клинических экспертов равнялся 7,6±7,5 года.

Таблица 3. Демография пациентов по стране														
Показатель	Всего (n=1806)	Бразилия (n=100)	Канада (n=53)	Китай (n=203)	Индия (n=209)	Италия (n=100)	Япония (n=168)	Ливан (n=103)	Мексика (n=153)	Нигерия (n=132)	Россия (n=104)	Южная Африка (n=208)	Испания (n=70)	Тунис (n=203)
Возраст, годы (среднее±СО)	39,9±13,7	32,9±9,6	39,8±14,2	43,9±15,6	36,5±11,4	41,4±11,2	47,0±15,1	36,4±12,5	38,1±13,0	37,5±12,2	36,3±11,7	35,1±11,0	52,0±16,2	43,2±12,6
Пол, n (%):														
мужской	908 (50,3)	62 (62,0)	19 (35,8)	123 (60,6)	120 (57,4)	50 (50,0)	72 (42,9)	38 (36,9)	48 (31,4)	65 (49,2)	44 (42,3)	133 (63,9)	26 (37,1)	108 (53,2)
женский	897 (49,7)	38 (38,0)	33 (62,3)	80 (39,4)	89 (42,6)	50 (50,0)	96 (57,1)	65 (63,1)	105 (68,6)	67 (50,8)	60 (57,7)	75 (36,1)	44 (62,9)	95 (46,8)
Семейное положение, n (%)														
холост/не замужем	992 (54,9)	81 (81,0)	22 (41,5)	110 (54,2)	66 (31,6)	71 (71,0)	77 (45,8)	68 (66,0)	91 (59,5)	68 (51,5)	65 (62,5)	167 (80,3)	28 (40,0)	78 (38,4)
в браке/сожительство	597 (33,1)	12 (12,0)	17 (32,1)	75 (36,9)	133 (63,6)	19 (19,0)	64 (38,1)	20 (19,4)	42 (27,5)	41 (31,1)	22 (21,2)	25 (12,0)	28 (40,0)	99 (48,8)
в разлуке/в разводе	177 (9,8)	6 (6,0)	13 (24,5)	15 (7,4)	4 (1,9)	7 (7,0)	21 (12,5)	15 (14,6)	20 (13,1)	18 (13,6)	13 (12,5)	13 (6,3)	9 (12,9)	23 (11,3)
вдовьяе	40 (2,2)	1 (1,0)	1 (1,9)	3 (1,5)	6 (2,9)	3 (3,0)	6 (3,6)	0	0	5 (3,8)	4 (3,8)	3 (1,4)	5 (7,1)	3 (1,5)
Занятость, n (%)														
полная	403 (22,3)	4 (4,0)	14 (26,4)	47 (23,2)	69 (33,0)	11 (11,0)	26 (15,5)	16 (15,5)	17 (11,1)	41 (31,1)	22 (21,2)	22 (10,6)	26 (37,1)	88 (43,3)
неполная	142 (7,9)	5 (5,0)	6 (11,3)	3 (1,5)	12 (5,7)	9 (9,0)	14 (8,3)	11 (10,7)	31 (20,3)	11 (8,3)	6 (5,8)	8 (3,8)	3 (4,3)	23 (11,3)
безработный	1009 (55,9)	76 (76,0)	30 (56,6)	80 (39,4)	110 (52,6)	74 (74,0)	109 (64,9)	66 (64,1)	79 (51,6)	64 (48,5)	53 (51,0)	167 (80,3)	20 (28,6)	81 (39,9)
обучающийся	136 (7,5)	6 (6,0)	4 (7,5)	15 (7,4)	15 (7,2)	4 (4,0)	10 (6,0)	15 (14,6)	30 (19,6)	10 (7,6)	7 (6,7)	12 (5,8)	2 (2,9)	6 (3,0)
пенсионер	152 (8,4)	10 (10,0)	1 (1,9)	62 (30,5)	3 (1,4)	2 (2,0)	15 (8,9)	0	5 (3,3)	8 (6,1)	18 (17,3)	0	22 (31,4)	6 (3,0)
Условия лечения, n (%)														
амбулаторные	801 (44,4)	82 (82,0)	53 (100)	0	122 (58,4)	67 (67,0)	48 (28,6)	14 (13,6)	135 (88,2)	84 (63,6)	4 (3,8)	0	49 (70,0)	143 (70,4)
стационарные	994 (55,0)	18 (18,0)	0	203 (100)	87 (41,6)	33 (33,0)	120 (71,4)	89 (86,4)	17 (11,1)	48 (36,4)	91 (87,5)	207 (99,5)	21 (30,0)	60 (29,6)
другое	11 (0,6)	0	0	0	0	0	0	0	1 (0,7)	0	9 (8,7)	1 (0,5)	0	0

Как показано в табл. 3, для исследования по Протоколу 1 и Протоколу 2 были привлечены 1806 пациентов – соответственно 1041 и 765. Средний возраст участвующих пациентов составлял 39,9±13,7 года, и возраст по странам был сопоставим. Общая выборка имела равное распределение по полу. Семейное положение большинства пациентов в общей выборке было следующим: одиноки – 54,9%, в браке/сожительство – 33,1%, разведены/расстались – 9,8%, и вдовьяе были 2,2%. Больше половины пациентов в общей выборке были безработными (55,9%), и только 22,3% пациентов имели полную занятость. Чуть больше половины привлеченных пациентов в общей выборке были стационарными (55,0%), почти все остальные – амбулаторными (44,4%). Небольшая оставшаяся доля пациентов (0,6%) были зарегистрированы в других типах программ, таких как дневной стационар.

### Сбор и обработка данных, администрирование

Докладываемые клиническими интервьюерами данные были собраны с помощью Electronic Field Study System (EFSS) – сетевой системы сбора данных, разработанной с использованием Qualtrics™ (Provo, UT, USA) и переведенной на пять языков, применяемых в исследовании. Клиницисты заходили в EFSS, используя уникальный пароль, и сообщали всю информацию по исследованию.

Данные из центров хранились и управлялись централизованно в центре координации данных (ЦКД) в Columbia University. Качество данных устанавливалось за счет продолжительного отслеживания процедуры сбора данных местными научными сотрудниками с использованием программных функций Qualtrics™, таких как опции принудительного ответа и проверки информационного наполнения. Это обеспечило механизм сбора данных из всех центров в стандартизованном единообразном формате. Исследовательские команды на местах хранили записи любых ошибок ввода, которые были направлены в ЦКД для коррекции.

### Анализ данных

Основной анализ в этом исследовании касался надежности диагнозов, включенных в Протоколы 1 и 2. Данные из обоих протоколов были объединены в данном анализе. Надежность диагноза рассчитывали, основываясь на соглашении между клиническими экспертами независимо от того, был ли диагноз первым, вторым или третьим в списке. Например, если для конкретного пациента один клинический эксперт выставил диагноз единичного эпизода депрессивного расстройства, паническое расстройство и агорафобию, а другой клинический эксперт диагностировал агорафобию и единичный эпизод депрессивного расстройства, то оба клинических эксперта будут считаться согласными по

Диагноз	Число диагнозов	Согласие совместных экспертов (межклассовая каппа)	Стандартная ошибка	95% бутстрэповский ДИ
Шизофрения	725	0,87	0,012	0,84–0,89
Шизоаффективное расстройство	189	0,66	0,035	0,58–0,72
Острое и транзиторное психотическое расстройство	40	0,45	0,087	0,27–0,60
Бредовое расстройство	30	0,69	0,084	0,51–0,84
Биполярное аффективное расстройство первого типа	351	0,84	0,017	0,81–0,87
Биполярное аффективное расстройство второго типа	95	0,62	0,048	0,52–0,70
Единичный эпизод депрессивного расстройства	191	0,64	0,035	0,57–0,77
Рекуррентное депрессивное расстройство	267	0,74	0,025	0,69–0,79
Дистимическое расстройство	57	0,45	0,073	0,28–0,58
Генерализованное тревожное расстройство	129	0,62	0,044	0,53–0,70
Паническое расстройство	59	0,57	0,069	0,42–0,69
Агорафобия	46	0,62	0,072	0,47–0,75
Социофобия	38	0,88	0,045	0,78–0,95
Посттравматическое стрессовое расстройство	51	0,49	0,076	0,33–0,64
Сложное посттравматическое стрессовое расстройство	45	0,56	0,077	0,40–0,71
Расстройство адаптации	82	0,73	0,046	0,63–0,81

диагнозам единичного эпизода депрессивного расстройства и агорафобии, но не согласными по паническому расстройству.

В этот анализ были включены только те диагнозы, которые ставились как минимум 30 раз за время исследования, поскольку диагнозы, которые встречались менее часто, рассматривались как не имеющие достаточной стабильности для осуществления экспертизы. Чтобы оценить надежность диагноза были использованы каппа-коэффициенты межклассовой надежности для диагнозов, взвешенных по центрам, а также в исследовании была рассчитана распространенность. Затем был рассчитан 95% бутстрэповский доверительный интервал (ДИ) для каппы, основываясь на 1000 образцов повторной выборки. Весь анализ проводился с использованием SPSS.

Показатели Landis и Koch<sup>38</sup> были использованы для описания диапазона значений надежности для каппа-коэффициента: от незначительной (от 0 до 0,20), удовлетворительной (от 0,21 до 0,40), умеренной (от 0,41 до 0,60), значительной (от 0,61 до 0,80) до почти идеальной (от 0,81 до 1,0).

## РЕЗУЛЬТАТЫ

Оценка согласия между экспертами представлена в табл. 4, наряду с 95% бутстрэповским ДИ. Точечная оценка каппы была в диапазоне от 0,45 (дистимическое расстройство) до 0,88 (социофобия) и считалась от умеренной до почти идеальной для всех диагнозов, для которых была рассчитана, согласно Landis и Koch.

Оценка каппы была почти идеальной для шизофрении (0,87) и биполярного расстройства первого типа (0,84); значительной – для шизоаффективного расстройства (0,66), бредового расстройства (0,69), биполярного расстройства второго типа (0,62), единичного эпизода депрессивного расстройства (0,64), рекуррентной депрессии (0,74), генерализованного тревожного расстройства (0,62), агорафобии (0,62) и расстройства адаптации (0,73); умеренной – для острого и транзиторного психотического расстройства (0,45), дистимического расстройства (0,45), панического расстройства (0,57), посттравматического стрессового расстройства (0,49) и для недавно представленного диагноза сложного посттравматического стрессового расстройства (0,56).

В общем, точечная оценка каппы была ниже для расстройств, для которых была получена меньшая выборка. Более высокое число диагнозов первичного психотического и аффективного расстройств отражает тип условий (55% стационарных пациентов) и суть центров (третичная и вторичная медицинская помощь), вовлеченных в ПИВС.

Оценка каппы была точной для всех диагнозов, для которых она была рассчитана (ДИ <0,5; стандартная ошибка <0,1). Низкая граница оценки ДИ для каппы была >0,4 (удовлетворительная надежность) для 13 из 16 расстройств. Тем не менее нижняя граница оценки была в удовлетворительном диапазоне (от 0,2 до 0,4) только для острого и транзиторного психотического расстройства (0,27), дистимического расстройства (0,28) и посттравматического стрессового расстройства (0,33). Все диагнозы с нижней границей ДИ оценки каппа (<0,4) выставлялись реже, и это может наводить на мысль, что надежность для этих расстройств может увеличиться в выборках большего размера.

Табл. 5 показывает сравнение результатов согласия связанных экспертов в данном исследовании КОДУ к МКБ-11 с результатами полевых исследований КОДУ к МКБ-10<sup>16</sup>. Это сравнение было задумано как более демонстративное, нежели точное, из-за огромной разницы в методологии исследований. В отличие от КОДУ к МКБ-11, где были задействованы два эксперта в совместном интервью лицом к лицу, в полевых исследованиях МКБ-10 применялись разборы клинических случаев, в которых один эксперт проводил интервью лицом к лицу, а затем представлял случай другим экспертам, что и служило основой для исследования межэкспертной надежности. Разбор клинических случаев как метод чаще приводит к более консенсусным результатам, в которых надежность может быть соответственно выше. Более того, хотя большая часть диагнозов МКБ-10 представляли собой нечто очень похожее на предложенное в МКБ-11, они не являются идентичными.

Хотя статистическое сравнение этих двух исследований несправедливо, в 10 из 14 возможных сравнений между ПИВС КОДУ к МКБ-11 и полевых исследованиях КОДУ к МКБ-10 значения каппа были выше для МКБ-11. Эти различия тяготеют к тому, чтобы быть умеренными.

Таблица 5. Сравнение оценки надежности в ПИВС КОДУ к МКБ-11 и полевых исследованиях КОДУ к МКБ-10.

ПИВС МКБ-11		Полевые исследования КОДУ к МКБ-10	
	Каппа (n)		Каппа (n)
Шизофрения	0,87 (725)	F20 Шизофрения	0,81 (490)
Шизоаффективное расстройство	0,66 (189)	F36 Шизоаффективное расстройство	0,48 (148)
Острое и транзиторное психотическое расстройство	0,45 (40)	F23 Острое и транзиторное психотическое расстройство	0,65 (146)
Бредовое расстройство	0,69 (30)	F22.0 Бредовое расстройство	0,62 (83)
Биполярное аффективное расстройство первого типа	0,84 (351)	F30 Маниакальный эпизод	0,69 (53)
		F31 Биполярные аффективные расстройства	0,81 (259)
Единичный эпизод депрессивного расстройства	0,64 (191)	F32 Депрессивный эпизод	0,66 (353)
Рекуррентное депрессивное расстройство	0,74 (267)	F33 Рекуррентное депрессивное расстройство	0,69 (302)
Дистимическое расстройство	0,45 (57)	F34.1 Дистимия	0,36 (101)
Генерализованное тревожное расстройство	0,62 (129)	F41.1 Генерализованное тревожное расстройство	0,48 (67)
Паническое расстройство	0,57 (59)	F41.0 Паническое расстройство	0,74 (31)
Агорафобия	0,62 (46)	F40.0 Агорафобия	0,51 (22)
Социофобия	0,88 (38)	F40.1 Социофобия	0,41 (22)
Посттравматическое стрессовое расстройство	0,49 (51)	F43.1 Посттравматическое стрессовое расстройство	0,62 (23)
Расстройство адаптации	0,73 (82)	F43.2 Расстройство адаптации	0,54 (107)

## ОБСУЖДЕНИЕ

11-й пересмотр глав МКБ о психических, поведенческих расстройствах и расстройствах психического развития предусматривает внесение значительных изменений в формирование концептуального представления о многих расстройствах, которые могут влиять на их надежность, валидность и клиническую полезность. Полевые исследования, изучающие, как хорошо предложенные изменения работают в руках пользователей, для которых они предназначены, были ключевыми в процессе пересмотра. Соответственно, ПИВС для КОДУ к МКБ-11 были проведены в широком спектре учреждений вторичной и третичной психиатрической помощи в странах с различными языками, культурой и уровнем ресурсов.

Результаты ПИВС к МКБ-11 показали, что все частые и несущие тяжелое бремя расстройства во взрослой популяции, рассматриваемые в данном исследовании, были диагностированы с как минимум удовлетворительной, а в большинстве случаев – отличной надежностью группой клинических экспертов, включавшей в себя продвинутых ординаторов по специальности «психиатрия» наряду с более опытными клиницистами. Это наводит на мысль, что предлагаемые КОДУ к МКБ-11 подходят для применения на уровне всего мира и что их успешное внедрение распространится с применения к письменным клиническим описаниям до применения на реальных пациентах в клинических условиях.

Надежность диагнозов влияет на клиническое взаимодействие, обобщаемость руководств на всю популяцию пациентов и подгонку лечения согласно диагнозу, а кроме того – на формирование выборки пациентов для исследования. DSM-III предложило полностью операциональные диагностические критерии в классификации психических расстройств для улучшения диагностической надежности<sup>30,31</sup>. КОДУ к МКБ-11 были созданы, чтобы привести ее в соответствие с всеохватывающими принципами категоризации, восходящие к ранним исследованиям, анализировавшим,

как клиницисты естественно организуют клинические состояния<sup>2</sup>. Расстройства МКБ-11 представлены в терминах основных характерных признаков, которые клиницисты имеют основания ожидать встретить во всех случаях, пытаясь обсудить сущность расстройства, имея большую гибкость к клиническим и культурным различиям<sup>13</sup>. КОДУ к МКБ-11 избегают полностью операциональных критериев, которые можно охарактеризовать четкими границами и подсчетом симптомов, кроме того, это поддерживается клиническим опытом. Настоящие результаты бросают вызов предположениям о том, что чем более дружественный клинику, менее алгоритмизированный и менее строго описанный подход используется для КУДО к МКБ-11, тем он менее надежен.

Коэффициенты надежности, наблюдаемые в данном исследовании, были основаны на обычных клинических обследованиях (длящихся около 1 часа) и использовании опроса в открытой форме клиницистами с различными подготовкой и опытом. Результаты были схожими с теми, что были получены с помощью диагностических обследований с использованием более сложных и времязатратных структурированных инструментов<sup>26,39,40</sup>. Эти результаты дают возможность предположить, что использование клиницистами однотипных процедур, освоенных в ходе короткого обучения, может давать достаточную надежность для часто диагностируемых психических расстройств в клинических условиях. Гипотеза, которую было бы неплохо протестировать, используя ресурсы, которые предназначены для совершенствования диагностических критериев, – это то, что дальнейшие успехи могли бы быть достигнуты за счет большей концентрации внимания на надлежащем обучении диагностическим умениям и техникам опроса<sup>41</sup>, нежели на достижении большей точности в строгих операциональных диагностических руководствах.

В целом надежность диагнозов в КОДУ к МКБ-11 была выше, чем у диагнозов в КОДУ к МКБ-10<sup>16</sup>, хотя строгое



сравнение неуместно из-за большой разницы в методологии этих полевых исследований. Похожим образом не проводилось сравнения исследовательских диагностических критериев МКБ-10 и DSM-III<sup>32</sup>, потому что методологические различия между ними еще сильнее. Полевое изучение исследовательских критериев МКБ-10 включало использование структурированных диагностических инструментов, которые охватывали диагностические критерии во время обследования<sup>33</sup>. Опубликованные результаты полевых исследований DSM-III дали значение каппа для групп расстройств больше, чем для специфических расстройств<sup>32</sup>, и это имеет склонность максимизировать результаты надежности, потому что различия внутри группы значительно более вероятны, чем различия, касающиеся расстройств из разных групп.

Изменения в КОДУ к МКБ-11 относительно КОДУ к МКБ-10 были предложены экспертами рабочих групп, основываясь на доступных научных сведениях и с исключительным вниманием к дополнительным источникам информации, связанным с клинической полезностью и всемирной применимостью. Ни в коем случае изменения не были направлены только на повышение надежности; кроме того, этой цели, вероятно, способствовало более целостное предоставление информации в КОДУ к МКБ-11, если сравнивать с КОДУ к МКБ-10<sup>13</sup>. Тем не менее вопрос, было ли следствием этих изменений общее снижение надежности КОДУ к МКБ-11 относительно КОДУ к МКБ-10, останется поводом для беспокойства.

Надежность генерализованного тревожного расстройства, агорафобии, социофобии и расстройства адаптации в КОДУ к МКБ-11 повысилась относительно КОДУ к МКБ-10. Это обнадеживает, потому что надежность более мягких расстройств относительно более тяжелых (например, шизофрении и биполярного аффективного расстройства) была ниже, чем в полевых исследованиях МКБ-10<sup>16,33</sup>. Данные, полученные в полевых исследованиях DSM-5, наводят на мысль, что расстройства, которые описываются скорее в общих чертах, имеют большую надежность<sup>25</sup>. Некоторое количество иерархических правил исключения было убрано для тревожных и связанных со стрессом расстройств в КОДУ к МКБ-11, потому что для них было недостаточно эмпирического подкрепления<sup>7</sup>. Схожим образом подтипы расстройства адаптации были убраны из КОДУ к МКБ-11, потому что для них было недостаточно доказательств их валидности или клинической полезности<sup>7</sup>.

Концептуализация генерализованного тревожного расстройства расширилась в КОДУ к МКБ-11, включив беспокойство в качестве альтернативной сущностной особенности генерализованной тревоги и сопровождающих ее физиологических симптомов<sup>7</sup>, что основано, в частности, на исследованиях, которые показывают, что беспокойство – главная характеристика данного расстройства<sup>42</sup>. Агорафобия была реконцептуализирована, чтобы включить более широкий набор устрашающих стимулов (страх ситуаций, страх особых негативных результатов) и поведение, появляющееся в ответ на эти стимулы (избегание или погружение в ситуации при особых условиях или претерпевание ситуации с интенсивным страхом/тревогой), в частности, чтобы допустить больше ситуаций, которые могут быть более репрезентативны в странах с низким или средним уровнем доходов<sup>43</sup>. Осмысление социофобии в МКБ-11 расширило ее суть в сторону, в которой человек может бояться быть негативно оцененным другими людьми, добавив культуральные различия этого расстройства (т. е. страх унижения, отказа или боязнь показаться грубым) наряду с рядом способов реагирования в ответ на социальные стимулы<sup>44,45</sup>. Возможно, что большее внимание к когнитивным и поведенческим компонентам тревоги и их контекстуальные и

поведенческие особенности в КОДУ к МКБ-11 относительно КОДУ к МКБ-10<sup>7</sup> помогли повысить надежность этих диагнозов.

Изменения, внесенные в диагностические руководства для расстройства адаптации, основаны на ранних исследованиях (типа случай-контроль) связанных со стрессом расстройств, в частности за счет обеспечения дополнительных указаний по отличиям от нормальных реакций на стресс, возможно, улучшило его диагностическую надежность в данном исследовании.

Шизоаффективное расстройство не является редким диагнозом в популяции больных, а его надежность стала субъектом проходящих в данное время дискуссий<sup>46</sup>. Jager и соавт.<sup>47</sup> пересмотрели шесть исследований и сообщили о размере каппы между 0,08 и 0,63, подытоживая тем самым, что лишь одно исследование показало хороший уровень согласия. В метаанализе исследований последовательной надежности (тест-ретест) шизоаффективного расстройства Santelmann и соавт.<sup>46</sup> задокументировали среднюю разницу каппы примерно в 0,2 между шизоаффективным расстройством и другими диагнозами, такими как шизофрения, биполярное расстройство и униполярная депрессия. Повысившаяся надежность шизоаффективного расстройства в МКБ-11 по сравнению с КОДУ к МКБ-10 может быть связана с решением, принятым в предлагаемых КОДУ к МКБ-11: применять диагностические требования еще более четко к текущему эпизоду, нежели к продолжительному курсу этой болезни<sup>3</sup>. Это отличается от основанного на выяснении длительности заболевания подхода исторически и в настоящее время применяемого в DSM, на чем основывалась большая часть последних исследований.

Некоторые области этой классификации заслуживают дальнейшего рассмотрения, основанного на полученных результатах. Диагнозы острого и транзиторного психотического расстройства, панического расстройства и посттравматического стрессового расстройства в КОДУ к МКБ-11, как оказалось, имели меньшую надежность, чем аналогичные категории в КОДУ к МКБ-10, хотя было решено, что неуместно статистически сравнивать эти различия. Тем не менее эти различия имеют довольно скромный размер (<0,2 во всех случаях), а оцениваемая надежность этих диагностических категорий в КОДУ к МКБ-11 все еще находится в умеренном диапазоне.

Следует отметить, что, в отличие от категорий, обсуждаемых ранее, которые были расширены в КОДУ к МКБ-11, описание каждой из этих категорий сузилось с точки зрения их сущностных особенностей. Острое и транзиторное психотическое расстройство в МКБ-11 теперь включает в себя исключительно острые психозы с демонстрируемым «полиморфизмом»<sup>3</sup>, который не может быть серьезно сравним с более обширной концепцией, апробированной в полевых исследованиях МКБ-10<sup>16</sup>. Надежность острого и транзиторного психотического расстройства с полиморфными симптомами в полевом исследовании исследовательских критериев к МКБ-10<sup>33</sup> была похожа на то, что было представлено в настоящем изучении. Все равно на основании этих результатов описание острого и транзиторного психотического расстройства было пересмотрено для финальной версии этого руководства, чтобы определить этот аспект расстройства более исключительно и обеспечить дополнительные указания о том, как дифференцировать его от других состояний.

Предлагаемые КОДУ к МКБ-11 для панического расстройства теперь требуют четкого разделения между паническими атаками непредсказуемой природы и паническими атаками, возникающими в связи с симптомами специфических психических расстройств (т. е. фобических расстройств, некоторых обсессивно-компульсивных расстройств и расстройств, связанных со стрессом). Если

паническую атаку можно представить как следствие симптомов другого специфического психического расстройства, то стоит применять спецификатор «с паническими атаками», нежели выставлять отдельно существующий диагноз панического расстройства. Если некоторые панические атаки в течение расстройства были неожиданными и не возникали обязательно в ответ на стимулы, обусловленные концентрацией внимания на тревоге, которая связана с соответствующим расстройством, стоит поставить отдельный диагноз панического расстройства. В таких случаях не нужно ставить спецификатор «с паническими атаками»<sup>7</sup>. Более низкое значение каппы для КОДУ к МКБ-11 относительно КОДУ к МКБ-10 для панического расстройства наводит на мысль, что клиницисты имели трудности в дифференциации ожидаемых и неожиданных панических атак или им было не ясно, когда использовать спецификатор «с паническими атаками», и они дополнительно использовали диагноз панического расстройства. Это представляет собой заметный пример компромисса между валидностью и надежностью. Основываясь на результатах данного исследования, финальная версия КОДУ к МКБ-11 содержит более детализированные указания о том, как дифференцировать неожиданные и ожидаемые панические атаки и как решать, применять ли спецификатор «с паническими атаками» или ставить дополнительный диагноз панического расстройства. Акцент на этой проблеме в образовательных программах может быть полезен в рамках внедрения МКБ-11.

Хотя посттравматическое стрессовое расстройство является хорошо распознаваемой клинической единицей, его критикуют за обширный состав кластера его симптомов и высокую частоту сочетаемости с другими расстройствами. В исследованиях также предполагалось, что порог для диагностики этого расстройства в МКБ-10 довольно низок<sup>48,49</sup>. В КОДУ к МКБ-11 диагноз посттравматического стрессового расстройства проще по своей сути, чем в МКБ-10, и теперь требует наличия повторного переживания навязчивых симптомов «здесь и сейчас», против только претерпевания навязчивых мыслей о травмирующем событии, наряду с наличием функциональных нарушений<sup>5</sup>. Эта модель набирает все усиливающееся эмпирическое обоснование<sup>50</sup>. Несмотря на то что более раннее интернет-исследование связанных со стрессом расстройств<sup>9</sup> показало, что клиницисты не всегда применяют предлагаемые к МКБ-11 инструкции, касающиеся необходимого элемента в виде повторного переживания травматического(их) события(ий). Поздняя версия КОДУ к МКБ-11, использованная в настоящем исследовании, способствовала внесению дополнительной ясности относительно повторного переживания при ПТСР. Тем не менее применение некоторых изменений для посттравматического стрессового расстройства, представленных в МКБ-11, до сих пор является сложным для практикующих клиницистов. Дальнейшее изучение этих противоречий между клиническими экспертами на уровне определенных симптомов поможет лучше понять эту проблему. Особое внимание к новой сущности посттравматического стрессового расстройства в рамках образовательных программ по МКБ-11 также очень желательно.

Некоторые ограничения ПИВС МКБ-11 нуждаются в озвучивании. Во-первых, может показаться повторением, но методика со связанными экспертами (одновременно) при проверке надежности, которая не дает информации, излагаемой двумя диагностами, быть идентичной, обычно повышает ценность индекса каппа относительно каппы, получаемой, когда обследования проводятся отдельно<sup>26,51</sup>. Во-вторых, настоящее исследование проводилось во множестве центров в разных странах, учитывая большую долю стран с низким и средним доходом, но участвовавшие клиницисты не могут считаться всемирно репрезентативной

выборкой профессионалов, диагностирующих психические расстройства. Участвующие учреждения обычно были высокостатусными центрами вторичной или третичной помощи, где обучение клиницистов диагностической классификации и обследованию, с большой вероятностью, соответствует самым высоким национальным стандартам. Также клинические интервьюеры, участвующие в исследовании, могли иметь особый интерес к диагностической классификации и к обучению работе с МКБ-11. Это можно трактовать как то, что надежность, полученная в данном исследовании, выше, чем та, что будет получена в рутинной практике во всех условиях, в которых КОДУ к МКБ-11 будет применена. Впрочем, эти проблемы присущи данной области исследований, разве что их можно улучшить уровнем ресурсов значительно более избыточным, чем тот, что был доступен для ПИВС.

Более того, поскольку исследовательские центры были большими академическими организациями, которые предназначены помогать пациентам с уровнем проблем с психическим здоровьем от умеренного до тяжелого, результаты могут быть не обобщаемыми на пациентов с более легкими расстройствами, которых можно встретить в условиях общества. Беспокойство на этот счет уменьшит тот факт, что КОДУ к МКБ-11 включает в себя специальные указания о разграничении расстройств от вариантов нормы и повысило диагностический порог для некоторых состояний, тестируемых в ПИВС (например, связанных со стрессом расстройств).

В итоге, настоящее исследование рассматривало лишь относительно малую часть широкого спектра диагнозов психических расстройств, которые могут применяться ко взрослым пациентам, обращая особое внимание на те, что ответственны за самый большой уровень бремени болезни, и те, что требуют большей части услуг в сфере психического здоровья в участвовавших центрах. Гораздо более широкий спектр диагностических категорий был рассмотрен с помощью интернет-исследований, и общая непротиворечивость между результатами этих двух типов исследований обнадеживает.

## ЗАКЛЮЧЕНИЕ

ПИВС КОДУ к МКБ-11 как экспериментальное полевое исследование<sup>36</sup> было разработано, чтобы получить информацию об источнике диагностических разночтений посредством изучения каждого элемента диагностического руководства для расстройств, включенных в определенный протокол. Это исследование дало дополнительные данные для ВОЗ, которые можно использовать для улучшения диагностических руководств до их публикации. ВОЗ также будет использовать данные для разработки обучающих руководств и обучающих занятий для клиницистов, чтобы помогать странам-участникам во внедрении МКБ-11, уделяя особое внимание странам с низким и средним доходом, в которых живет подавляющее большинство населения мира.

Основной вывод этого международного исследования в том, что предложенные КОДУ к МКБ-11 могут восприниматься профессионалами в области диагностики психического здоровья в широком диапазоне стран согласованно. Всемирная возможность применения часто диагностируемых психических расстройств КОДУ к МКБ-11 поддерживается оцененной надежностью этих руководств в различных условиях (в 28 центрах в 13 странах и на 5 языках), используя естественный дизайн полевого исследования и подход к обучению, который можно легко повторить при внедрении МКБ-11. За исключением ограниченного числа состояний, в которых мы потерпели неудачу, выводы исследования будут информировать последующие переосмотры вплоть до публикации МКБ-11.

Размах данного сотрудничества, включение в практику клиницистов по всему миру, управление обучением на многих языках и завершение этого исследования вовремя, что позволяет внести изменения в финальную версию руководства, – главные сильные стороны программы исследования МКБ-11. В дополнение к особой ценности этого исследования в образовании МКБ-11, ПИВС и всемирная сеть клинической практики для проводимых в интернете полевых исследований МКБ-11 ВОЗ (<http://gcr.network>) оживили интерес среди клиницистов по всему миру к участию в проводимом исследовании, и это позволит и далее улучшать многие измерения клинического понимания психических болезней и способов оказания психиатрической помощи.

#### Благодарности

Мнения, содержащиеся в данной работе, принадлежат их авторам, кроме тех, которые особо помечены, и не намерены представлять официальные правила или позицию Всемирной организации здравоохранения. Для национальной деятельности, связанной с этим проектом, финансирование получили следующие страны: Бразилия – Conselho Nacional de Desenvolvimento Científico e Tecnológico; Канада – University Medical Research Fund, Royal's University of Ottawa Institute of Mental Health Research; Япония – Japanese Society of Psychiatry and Neurology и Japan Agency for Medical Research and Development; Мексика – Consejo Nacional de Ciencia y Tecnología. Дополнительная помощь в сборе данных в Бразилии, Ливане, Нигерии, Южной Африке и Тунисе была оказана программой всемирного психического здоровья Columbia University. В противном случае этот проект финансировался добровольными пожертвованиями участвующих учреждений. Авторы выражают свою признательность следующим людям, которые внесли значительный вклад в проведение исследования: Gustavo M. Barros, Ary Gadelha, Michel Haddad, Nuno H.P. Santos (Бразилия); Huajian Ma, Zhen Wang, Jingjing Huang (Китай); Huma Kamal, Nidhi Malhotra (Индия); Gaia Sampogna, Lucia Del Gaudio, Giuseppe Piegari, Francesco Perris, Luca Steardo Jr (Италия); Tomofumi Miura, Itta Namamura, Kiyokazu Atake, Ayako Endo, Yuki Kako, Shinichi Kishi, Michihiko Koeda, Shinsuke Kondo, Akeo Kurumaji, Shusuke Numata, Naoya Oribe, Futoshi Suzuki, Masashi Yagi (Япония); Sariah Daouk, Chadia Haddad, François Kazour, Nicole Khauli (Ливан); Francisco Juárez, Alejandra González, Omar Hernández, Carolina Muñoz (Мексика); Mayokun Odunleye (Нигерия); Tatiana Kiska, Oleg Limankin, Pavel Ponizovsky (Российская Федерация); Roxanne James, Christine Lochner, Adele Pretorius (Южная Африка); Carolina Ávila, Cora Fernández, Julián Gómez, Ana Izquierdo, Beatriz Vicario, Rubén Vicente (Испания); Rahma Damak (Тунис).

#### Библиография

- International Advisory Group for the Revision of ICD-10 Mental and Behavioural Disorders. A conceptual framework for the revision of the ICD-10 classification of mental and behavioural disorders. *World Psychiatry* 2011;10:86-92.
- Reed GM, Roberts MC, Keeley J et al. Mental health professionals' natural taxonomies of mental disorders: implications for the clinical utility of the ICD-11 and the DSM-5. *J Clin Psychol* 2013;69:1191-212.
- Gaebel W. Status of psychotic disorders in ICD-11. *Schizophr Bull* 2012;38:895-8.
- Maj M, Reed GM. The ICD-11 classification of mood and anxiety disorders: background and options. *World Psychiatry* 2012;11(Suppl. 1).
- Maercker A, Brewin CR, Bryant RA et al. Diagnosis and classification of disorders specifically associated with stress: proposals for ICD-11. *World Psychiatry* 2013;12:198-206.
- Stein DJ, Kogan CS, Atmaca M et al. The classification of obsessive-compulsive and related disorders in the ICD-11. *J Affect Disord* 2016;190:663-74.
- Kogan CS, Stein DJ, Maj M et al. The classification of anxiety and fear-related disorders in the ICD-11. *Depress Anxiety* 2016;33:1141-54.
- Tyrer P, Reed GM, Crawford MJ. Classification, assessment, prevalence and effect of personality disorder. *Lancet* 2015;385:717-26.
- Keeley JW, Reed GM, Roberts MC et al. Disorders specifically associated with stress: a case-controlled field study for ICD-11 Mental and Behavioural Disorders. *Int J Clin Health Psychol* 2016;16:109-27.
- World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva: World Health Organization, 1992.
- Evans SC, Roberts MC, Keeley JW et al. Using vignette methodologies for study clinicians' decision-making: validity, utility, and application in ICD-11 field studies. *Int J Clin Health Psychol* 2015;15:160-70.
- Keeley JW, Reed GM, Roberts MC et al. Developing a science of clinical utility in diagnostic classification systems: field study strategies for ICD-11 mental and behavioural disorders. *Am Psychol* 2016;71:3-16.
- First MB, Reed GM, Hyman SE et al. The development of the ICD-11 Clinical Descriptions and Diagnostic Guidelines for Mental and Behavioural Disorders. *World Psychiatry* 2015;14:82-90.
- Reed GM. Toward ICD-11: improving the clinical utility of WHO's international classification of mental disorders. *Prof Psychol Res Pr* 2010;41:457-64.
- Reed GM, First MB, Medina-Mora ME et al. Draft diagnostic guidelines for ICD-11 mental and behavioural disorders available for review and comment. *World Psychiatry* 2016;15:112-3.
- Sartorius N, Kaelber CT, Cooper JE et al. Progress toward achieving a common language in psychiatry. Results from the field trial of the clinical guidelines accompanying the WHO classification of mental and behavioral disorders in ICD-10. *Arch Gen Psychiatry* 1993;50:115-24.
- Mullins-Sweatt SN, Widiger TA. Clinical utility and DSM-5. *Psychol Assess* 2009;21:302-12.
- Blashfield RK, Keeley JW, Flanagan EH et al. The cycle of classification: DSM-I through DSM-5. *Annu Rev Clin Psychol* 2014;10:25-51.
- Stengel E. Classification of mental disorders. *Bull World Health Organ* 1959;21:601-63.
- Spitzer R, Sheehy M, Endicott J. DSM-III: Guiding principles. In: Rakoff V, Stancer H, Kedward H (eds). *Psychiatric diagnosis*. New York: Brunner/Mazel, 1977:1-24.
- World Health Organization. Glossary of mental disorders and guide to their classification. Geneva: World Health Organization, 1974.
- Beck AT, Ward CH, Mendelson M et al. Reliability of psychiatric diagnosis 2: a study of consistency of clinical judgments and ratings. *Am J Psychiatry* 1962;119:351-7.
- Ward CH, Beck AT, Mendelson M et al. The psychiatric nomenclature: reasons for diagnostic disagreement. *Arch Gen Psychiatry* 1962;7:198-205.
- Harvey PD, Heaton RK, Carpenter WT Jr et al. Diagnosis of schizophrenia: consistency across information sources and stability of the condition. *Schizophr Res* 2012;140:9-14.
- Regier DA, Narrow WE, Clarke DE et al. DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnosis. *Am J Psychiatry* 2013;170:59-70.
- Chmielewski M, Clark LA, Bagby RM et al. Method matters: understanding diagnostic reliability in DSM-IV and DSM-5. *J Abnorm Psychol* 2015;124:764-9.
- Lahey BB, Applegate B, Barkley RA et al. DSM-IV field trials for oppositional defiant disorder and conduct disorder in children and adolescents. *Am J Psychiatry* 1994;151:1163-71.
- Rounsaville BJ, Kosten TR, Williams JB et al. A field trial of DSM-III-R psychoactive substance dependence disorders. *Am J Psychiatry* 1987;144:351-5.
- Brittain PJ, Stahl D, Rucker J et al. A review of the reliability and validity of OPCRIT in relation to its use for the routine clinical assessment of mental health patients. *Int J Methods Psychiatr Res* 2013;22:110-37.
- Aboraya A, Rankin E, France C et al. The reliability of psychiatric diagnosis revisited: the clinician's guide to improve the reliability of psychiatric diagnosis. *Psychiatry* 2006;3:41-50.
- First MB. Counterpoint – there isn't enough evidence available to speculate on the reliability of diagnoses in clinical settings. *Psychiatry* 2007;4:24-5.
- Spitzer R, Forman J, Nee J. DSM-III field trials, I: initial interrater diagnostic reliability. *Am J Psychiatry* 1979;136:815-7.

33. Sartorius N, Ustün TB, Korten A et al. Progress toward achieving a common language in psychiatry, II: Results from the international field trials of the ICD-10 diagnostic criteria for research for mental and behavioral disorders. *Am J Psychiatry* 1995;152:1427-37.
34. Clarke DE, Narrow WE, Regier DA et al. DSM-5 field trials in the United States and Canada, Part I: study design, sampling strategy, implementation, and analytic approaches. *Am J Psychiatry* 2013;170:43-58.
35. Feinn R, Gelernter J, Cubells JF et al. Sources of unreliability in the diagnosis of substance dependence. *J Stud Alcohol Drugs* 2009;70:475-81.
36. First MB. The importance of developmental field trials in the revision of psychiatric classifications. *Lancet Psychiatry* 2016;3:579-84.
37. Whiteford HA, Degenhardt L, Rehm J et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* 2013;382:1575-86.
38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
39. Andrews G, Peters L, Guzman A-M et al. A comparison of two structured diagnostic interviews: CIDI and SCAN. *Aust N Z J Psychiatry* 1995;29:124-32.
40. Lobbstaël J, Leurgans M, Arntz A. Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). *Clin Psychol Psychother* 2011;18:75-9.
41. Bruehl S, Ohrbach R, Sharma S et al. Approaches to demonstrating the reliability and validity of core diagnostic criteria for chronic pain. *J Pain* 2016;17(Suppl. 9):T118-31.
42. Olatunji BO, Broman-Fulks JJ, Bergman SM et al. A taxometric investigation of the latent structure of worry: dimensionality and associations with depression, anxiety, and stress. *Behav Ther* 2010;41:212-28.
43. Stein DJ. Agoraphobia and panic disorder: options for ICD-11. *World Psychiatry* 2012;11(Suppl. 1):89-93.
44. Emmelkamp PMG. Specific and social phobias in ICD-11. *World Psychiatry* 2012;11(Suppl. 1):94-9.
45. Kinoshita Y, Chen J, Rapee RM et al. Cross-cultural study of conviction subtype Taijin Kyofu: proposal and reliability of Nagoya-Osaka diagnostic criteria for social anxiety disorder. *J Nerv Ment Dis* 2008;196:307-13.
46. Santelmann H, Franklin J, Bußhoff J, et al. Test-retest reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression – a systematic review and metaanalysis. *Bipolar Disord* 2015;17:753-68.
47. Jager M, Haack S, Becker T et al. Schizoaffective disorder – an ongoing challenge for psychiatric nosology. *Eur Psychiatry* 2011;26:159-65.
48. Brewin CR, Fuchkan N, Huntley Z et al. Outreach and screening following the 2005 London bombings: usage and outcomes. *Psychol Med* 2010;40:2049-57.
49. Peters L, Slade T, Andrews G. A comparison of ICD-10 and DSM-IV criteria for posttraumatic stress disorder. *J Trauma Stress* 1999;12:335-43.
50. Brewin CR, Cloitre M, Hyland P et al. A review of current evidence regarding the ICD-11 proposals for diagnosing PTSD and complex PTSD. *Clin Psychol Rev* 2017;58:1-15.
51. Kraemer HC. The reliability of clinical diagnoses: state of the art. *Annu Rev Clin Psychol* 2014;10:111-30.
52. Reed GM, Rebello TJ, Pike KM et al. WHO's Global Clinical Practice Network for mental health. *Lancet Psychiatry* 2015;2:379-80.

DOI:10.1002/wps.20524