

Тяжесть психических расстройств

Mark Zimmerman, Theresa A. Morgan, Kasey Stanton

Department of Psychiatry and Human Behavior, Brown University School of Medicine, Rhode Island Hospital, Providence, RI, USA

Перевод: Мурашко А.А. (Москва)

Редактура: к.м.н. Рукавишников Г.В. (Санкт-Петербург)

Вопрос об оценке тяжести психических расстройств имеет важное клиническое значение. Так, степень тяжести влияет на решения об объеме необходимой медицинской помощи и на решения о социальной помощи вследствие инвалидности по психическому расстройству. Противоречия существуют как по вопросам эффективности антидепрессантов в зависимости от степени тяжести депрессии, так и в том, что касается предпочтительности медикаментозных методов перед психотерапией при тяжелой депрессии. Степени тяжести используются для оценки исходов в исследованиях терапии и могут применяться в качестве предикторов в клинической практике. Но что на самом деле означает сказать, что у кого-то тяжелое заболевание? Означает ли тяжесть количество симптомов у пациента? Их выраженность? Частоту или персистенцию? Влияние симптомов на функционирование или качество жизни? Вероятность того, что заболевание приведет к бессрочной инвалидности или смерти? Не затрагивая вопрос операционализации тяжести расстройств, рассмотрим другой аспект – должна ли тяжесть концептуализоваться одинаково для всех заболеваний или ей следует быть специфичной для разных расстройств? В данной работе мы оцениваем, как характеризуют степени тяжести в исследованиях и современных психиатрических диагностических системах с акцентом на депрессию и расстройства личности. Наш обзор показывает, что в DSM-5 тяжесть различных расстройств оценивается по-разному и исследователи применяют множество способов определения тяжести как депрессии, так и расстройств личности, хотя тяжесть первой в основном определяется в соответствии со значениями шкал оценки симптомов, тогда как тяжесть вторых чаще определяется нарушениями функционирования. Ввиду того что влияние на функционирование симптоматических расстройств зависит от факторов, не связанных с собственно расстройством, таких как самооценочность, жизнестойкость, копинг-способности, социальная поддержка, культурные и социальные ожидания, а также обязанности, связанные с жизненной ролью, и готовность окружения выполнять данные обязанности, мы утверждаем, что тяжесть подобных расстройств должна определяться независимо от нарушений функционирования.

Ключевые слова: степень тяжести, психические расстройства, нарушения функционирования, симптомы, депрессия, расстройства личности, трансдиагностические модели, Ni-TOP, DSM-5, МКБ-10.

(World Psychiatry 2018;17(3):258-275)

Определение тяжести заболевания имеет важное клиническое значение. В зависимости от расстройства, степень тяжести влияет на решение о необходимости проведения лечения, виде, интенсивности и необходимости продолжения или прекращения терапии. Оценка тяжести также влияет на прогноз в отношении возможности исполнения повседневного ролевого функционирования и определения нетрудоспособности. Степень тяжести используется для оценки результатов в клинических исследованиях и может быть значимым предиктором для клинической практики.

Но что на самом деле означает сказать, что у кого-то тяжелое заболевание? Из множества определений слова «тяжелый» наиболее подходящим в контексте заболевания является «в большой степени». Это определение, однако, не передает того, что имеется в виду, когда болезнь называют «тяжелой». Означает ли тяжесть количество симптомов, имеющихся у пациента? Выраженность симптомов? Их частоту или персистенцию? Влияние симптомов на функционирование или качество жизни? Вероятность того, что заболевание может привести к бессрочной инвалидности или смерти?

Некоторые из этих вопросов можно рассмотреть подробнее. Например, по поводу вероятности летального исхода: связана ли тяжесть заболевания с неминуемой смертью, смертью в ближайшем будущем или смертью в любой период в будущем? Также должен ли учитываться фактор врачебного вмешательства? Иными словами, является ли болезнь тяжелой только тогда, когда смерть вероятна, если болезнь не лечить, или только тогда, когда смерть вероятна независимо от вмешательства?

Возможно, определение степени тяжести должно быть независимым от нарушений функционирования или прогноза и основываться на структурных или морфологических изменениях при повреждении пораженного органа. Конечно, это не относится ко многим заболеваниям, но,

когда это можно измерить, должно ли это быть руководящим принципом для оценки тяжести?

Отложив вопрос о том, как следует обеспечить практическое применение понятия тяжести, необходимо рассмотреть еще одно соображение – следует ли концептуализовать тяжесть одинаково для всех болезней или она должна быть специфичной для различных расстройств? Следует ли оценивать тяжесть сердечной недостаточности, ревматоидного артрита, диабета, острой инфекции верхних дыхательных путей и головной боли в соответствии с общим стандартом или метрикой, или же каждое расстройство должно иметь свои собственные соответствующие рекомендации по оценке тяжести?

В этой статье мы рассмотрим, как описывают тяжесть расстройств в психиатрических исследованиях и современных психиатрических диагностических системах. Чтобы проиллюстрировать некоторые из проблем и противоречий в определении тяжести психических расстройств, мы сделали акцент на депрессии и расстройствах личности (РЛ). Клиническое значение оценки тяжести депрессии отражено в официальных рекомендациях по лечению, которые основаны на тяжести заболевания^{1,2}. Важность оценки тяжести РЛ отражена в предложении МКБ-11 о замене специфичных критериев для различных расстройств одной категорией расстройства личности, которая классифицируется по степени тяжести^{3,4}.

Прежде чем обсуждать вопрос о тяжести психических расстройств, мы представляем краткий обзор того, как тяжесть заболеваний концептуализируется, оценивается и измеряется при различных соматических заболеваниях, подчеркивая разнообразие подходов.

ТЯЖЕСТЬ СОМАТИЧЕСКИХ ЗАБОЛЕВАНИЙ

Не существует единого мнения или общего принципа для разграничения соматических заболеваний по степени тяже-

сти. В некоторых случаях тяжесть определяется степенью структурного повреждения пораженного органа. Например, тяжесть ревматоидного артрита определяется по рентгенологическим признакам поражения сустава⁵. Тяжесть диалитической ретинопатии оценивается в соответствии со степенью повреждения сетчатки, оцененной при непосредственном офтальмологическом обследовании⁶. Аналогичным образом, для определения тяжести некоторых болезней используются физиологические показатели, отражающие влияние заболевания на орган. Например, фракция выброса левого желудочка используется в качестве показателя тяжести сердечно-сосудистых заболеваний⁷⁻¹⁰. В качестве показателя тяжести муковисцидоза используется объем форсированного выдоха¹¹. Для оценки тяжести гепатита используют уровни аминотрансфераз и билирубина¹².

Иногда тяжесть заболевания определяется клиническим обследованием. Например, степень тяжести ревматоидного артрита оценивается не только с помощью рентгенологического обследования, но и в соответствии с учетом числа воспаленных и болезненных суставов¹³.

Тяжесть заболевания также определяется более широко, чтобы охватить показатели функционирования пораженного органа и связанные с этим эффекты. В исследовании прогностических последствий тяжести заболевания после сердечного приступа показатели тяжести были основаны на сердечно-легочной дисфункции и неврологическом статусе^{14,15}. Тяжесть серповидно-клеточной анемии была основана на наличии и частоте таких осложнений, как почечная недостаточность, некроз бедренных и плечевых костей и желчнокаменная болезнь¹⁶. В исследованиях тяжести хронической обструктивной болезни легких индекс BODE (B – индекс массы тела; O – обструкция дыхательных путей, измеренная объемом форсированного выдоха за одну секунду; D – шкала одышки; E – способность к физической нагрузке, измеренная тестом шестиминутной ходьбы) включает в себя не только степень повреждения легких. Было установлено, что индекс в целом является лучшим предиктором смертности, госпитализации в будущем, качества жизни и депрессии, чем только объем форсированного выдоха¹⁷. Унифицированная рейтинговая шкала болезни Паркинсона содержит четыре подшкалы, оценивающие психическое состояние, жизнедеятельность, двигательные нарушения и осложнения заболевания^{18,19}.

Отходя от прямой, или физиологической, оценки пораженного органа, Функциональная классификация Нью-Йоркской ассоциации кардиологов (New York Heart Association – NYHA) представляет собой способ оценки тяжести сердечно-сосудистых заболеваний, основанный на ограничениях физической активности и наличии физических симптомов, связанных с различной степенью активности²⁰.

В отличие от специфичных для заболевания физических и физиологических показателей тяжести, существуют сводные показатели тяжести состояния, такие как Шкалы клинической оценки тяжести состояния пациента (APACHE) и Упрощенная шкала оценки физиологических расстройств (SAPS), основанные на неспецифических клинических и биологических показателях состояния здоровья, таких как температура тела, возраст, полиорганная недостаточность в анамнезе, электролитный баланс и гематокрит^{21,22}. Эти показатели оценки тяжести заболевания были использованы для прогнозирования смертности в выборках как с гетерогенными, так и с одинаковыми расстройствами у пациентов в остром состоянии в отделении неотложной помощи и у госпитализированных пациентов^{23,24}.

Наконец, были разработаны самоопросники с целью оценки тяжести некоторых соматических заболеваний. Тяжесть доброкачественной гипертрофии предстательной железы оценивается по индексу симптомов Американской урологической ассоциации, основанному на частоте симп-

томов²⁵. Индекс тяжести тиннитуса основан на частоте функциональных нарушений или психологических симптомов, связанных с тиннитусом²⁶. Шкала оценки тяжести симптомов при синдроме раздраженного кишечника оценивает частоту, дистресс и нетрудоспособность в связи с заболеванием²⁷. Тяжесть головных болей оценивается при помощи Опросника по оценке влияния головных болей по частоте, средней интенсивности болей и связанным с ним нарушениям функционирования²⁸. Ливерпульская шкала тяжести приступов оценивает восприятие контроля приступа и тяжесть иктальных и постиктальных симптомов²⁹.

Clark и соавт.³⁰ объединили подход, принятый для разработки показателей самоотчета тяжести болезни для шести болезненных состояний в исследовании здоровья ветеранов. Они определяли тяжесть заболевания с точки зрения восприятия пациентами выраженности симптомов или осложнений заболевания, которые связаны со снижением качества жизни или состояния здоровья. Они различают тяжесть заболевания и влияние заболевания на функционирование (например, ухудшение состояния, удовлетворенность жизнью, благополучие), поскольку влияние на функционирование часто опосредовано личностными характеристиками (например, жизнестойкостью, самоэффективностью) и социальным контекстом.

ТЯЖЕСТЬ ПСИХИЧЕСКИХ РАССТРОЙСТВ В СООТВЕТСТВИИ С DSM-5

В отличие от некоторых соматических заболеваний, не существует специфических или неспецифических биомаркеров психических расстройств, которые достоверно характеризуют тяжесть расстройства. В отсутствие таких биологических или структурных показателей исследователям и клиницистам остается лишь оценивать эпифеномены психического расстройства, чтобы судить о его тяжести.

Вопросы распределения ресурсов в здравоохранении часто затрагивают пациентов с тяжелыми психическими заболеваниями, хотя не существует консенсуса в том, как определить такие состояния^{31,32}. В DSM-5³³, как и в предыдущих версиях классификации, тяжесть определяется только для некоторых расстройств. В таблице перечислены расстройства, для которых определена степень тяжести по DSM-5.

Подход DSM-5 к определению степени тяжести варьирует в зависимости от расстройства. Четыре степени тяжести умственной отсталости (легкая, умеренная, тяжелая, глубокая) являются наиболее детально описанными. При этом характеристики адаптивного функционального дефицита занимают по три страницы описания для каждой степени тяжести. В DSM-5 указано, что тяжесть была определена в соответствии с адаптивным уровнем функционирования, а не баллами IQ, потому что первый является лучшим показателем необходимости врачебной помощи. Аналогичным образом, уровень дефицита и выраженность функциональных нарушений, определяющих тяжесть расстройств аутистического спектра, связаны с потребностью в оказании помощи. Тяжесть расстройств обучения определяется как трудностями в приобретении учебных навыков, так и вероятностью приобретения этих навыков с вмешательством или без такового в будущем. Например, в DSM-5 тяжелое расстройство обучения определяется как «серьезные нарушения в приобретении навыков обучения, затрагивающие несколько академических областей, при которых человек вряд ли овладеет вышеуказанными навыками без постоянного интенсивного индивидуального и специализированного обучения в течение большей части школьных лет». Таким образом, для этих расстройств определение тяжести связано с предполагаемой степенью вмешательства.

Депрессия и мания классифицируются как легкие, умеренные или тяжелые в зависимости от количества симпто-

Характеристика тяжести расстройств в DSM-5	
Расстройство по DSM-5	Признаки, используемые для определения тяжести
Большое депрессивное расстройство (БДР)	Количество симптомов, степень дистресса, вызванного выраженностью симптомов, и нарушения в социальной и трудовой сферах
Мания, гипомания	Те же, что и при БДР
Расстройство употребления алкоголя	Число критериев
Расстройство употребления психоактивных веществ	Число критериев
Нервная булимия	Частота эпизодов компенсаторного поведения в неделю
Нервная анорексия	Индекс массы тела
Приступообразное переедание	Частота эпизодов переедания
Расстройства обучения	Тяжесть дефицита навыков и вероятность их приобретения с вмешательством или без него
Синдром дефицита внимания и гиперактивности	Количество симптомов, тяжесть индивидуальных симптомов или степень нарушения функционирования вследствие симптомов
Умственная отсталость	Уровень адаптации
Расстройства аутистического спектра	Степень дезадаптации, связанная с нарушениями вербальной и невербальной коммуникации, ригидность поведения, сложности повторения действий с изменением или ограниченным/повторяющееся поведение
Расстройство стереотипных движений	Легкость, с которой симптомы могут быть подавлены, и необходимость вмешательства для предотвращения серьезных травм
Психотические расстройства	Количественная оценка по 5-балльной шкале основных признаков психоза (бред, галлюцинации, дезорганизация речи, психомоторные нарушения и негативные симптомы). Оценка основана на выраженности симптомов или субъективном дистрессе вследствие симптома
Реактивное расстройство привязанности	Определена только тяжелая степень. Для установления тяжелой степени все критерии должны быть на высоком уровне
Расторможенное социальное расстройство	Определена только тяжелая степень. Для установления тяжелой степени все критерии должны быть на высоком уровне
Психосоматические расстройства	Количество критериев и соматические жалобы
Психологические факторы, влияющие на другие заболевания	Степень влияния на медицинское состояние или медицинский риск
Гиперсомния	Количество дней в неделю с трудностями в поддержании состояния бодрствования днем
Нарколепсия	Частота катаплексии и эффективность ответа на препараты, количество засыпаний в сутки, степень нарушения ночного сна
Обструктивное апноэ сна/гипопноэ	Показатель индекса апноэ/гипопноэ
Ночные кошмары	Частота кошмаров за неделю
Сексуальные дисфункции	Степень дистресса в связи с симптомами
Преждевременная эякуляция	Время до эякуляции
Сексуальная дисфункция, вызванная препаратом/веществом	Процент случаев, в которых указанная дисфункция появляется
Вызывающее оппозиционное расстройство	Количество ситуаций, в которых появляются симптомы
Поведенческое расстройство	Количество проблем с поведением или степень вреда, причиненного другим
Нейрокогнитивные расстройства	Степень ограничения способности к самообслуживанию

мов, уровня стресса вследствие выраженности симптомов и степени нарушения социального и профессионального функционирования. Тяжесть расстройств, связанных с употреблением алкоголя и психоактивных веществ, определяется на основе наличия ряда критериев (легкая – два или три критерия; умеренная – четыре или пять критериев; тяжелая – шесть или более критериев). Тяжесть синдрома дефицита внимания и гиперактивности зависит от количества симптомов, тяжести отдельных симптомов или степени нарушения функционирования вследствие присутствия симптомов. Тяжесть булимии определяется в соответствии с количеством неадекватных компенсаторных действий в неделю (легкая – 1–3; умеренная – 4–7; тяжелая – 8–13; экстремальная – 14 и более), хотя степень тяжести может быть выше при вовлечении других симптомов или уровня нарушений функционирования. Для нервной анорексии тяжесть определяется по индексу массы тела, а для приступообразного переедания – по количеству эпизодов переада-

ния в неделю, хотя, как и для нервной булимии, степень тяжести может быть увеличена при вовлечении других симптомов или уровня нарушений функционирования. Тяжесть сексуальных расстройств определяется уровнем дистресса, вызываемого симптомами. Исключение представляет преждевременная эякуляция, для которой тяжесть определяется временем до эякуляции. Тяжесть катаплексии частично определяется отсутствием реакции на медикаментозное лечение.

Этот краткий обзор иллюстрирует вариабельность подходов, принятых в DSM-5 для определения степени тяжести. В некоторых случаях акцент сделан на количестве представленных критериев, в некоторых – на основной особенности расстройства, в некоторых – на уровне дистресса, а в некоторых – на реакции на вмешательство и прогнозе. В отличие от многих соматических заболеваний, ни один из подходов к определению тяжести не связан с вероятностью неминуемой или будущей смерти, и большинство подходов

не связаны с прогнозом или дальнейшим течением расстройства. Большинство подходов к определению тяжести в DSM-5 связаны с числом симптомов или критериев расстройства, частотой симптомов и уровнем нарушений функционирования или дистресса.

ТЯЖЕСТЬ ДЕПРЕССИИ

Мы акцентируем внимание на тяжести депрессии, потому что по данному вопросу проведено наибольшее количество исследований. Несмотря на противоречия в исследованиях, тяжесть депрессии связана с качеством жизни³⁴, нарушениями функционирования³⁵⁻³⁶, суицидальным риском³⁷⁻³⁹, долгосрочным прогнозом⁴⁰⁻⁴³ и несколькими биологическими переменными⁴⁴⁻⁴⁶. Кроме того, тяжесть депрессии вызывает разногласия относительно эффективности лечения и того, следует ли рекомендовать определенные формы лечения в качестве терапии первой линии. Почти все исследования тяжести депрессии основаны на баллах оценочных шкал, однако большинство опросников были разработаны без учета того, как лучше всего концептуализировать и оценивать тяжесть депрессии.

Степени тяжести депрессии в DSM-5 и МКБ-10

Для определения степени тяжести депрессии в DSM-5 используются три критерия: количество симптомов, уровень дистресса вследствие выраженности симптоматики и уровень нарушения социального и профессионального функционирования. Категоризация тяжести депрессии применима ко всем расстройствам депрессивного спектра, не только к БДР. Легкая депрессия устанавливается при наличии «малого количества, если таковые имеются, симптомов сверх тех, которые необходимы для выставления диагноза, выраженность симптомов вызывает тревогу, но поддается контролю, а симптомы приводят к незначительным нарушениям социального или профессионального функционирования». Тяжелая депрессия устанавливается в тех случаях, когда «число имеющихся симптомов значительно превышает число симптомов, необходимое для постановки диагноза, уровень проявления симптомов вызывает выраженный дистресс и не поддается контролю, а симптомы значительно влияют на социальное и профессиональное функционирование». Умеренная депрессия четко не определена в DSM-5. Указывается, что число симптомов, их выраженность и/или нарушения функционирования определяются между легкой и тяжелой степенью.

В DSM-5 есть ряд затруднений, связанных с описанием степеней тяжести. Так, одно и то же определение спецификатора тяжести используется и для БДР, и для дистимии. Это является проблемой, потому что для диагностики дистимии в DSM-5 требуется меньшее количество симптомов, чем для БДР. Таким образом, пациент с дистимией, имеющих то же количество симптомов, что и пациент с БДР, и с аналогичным уровнем функциональных нарушений и дистресса может быть расценен как более тяжелый, поскольку количество симптомов может «существенно превышать» диагностический порог для дистимии, но не для БДР.

Другая проблема спецификатора тяжести DSM-5 заключается в том, что определение нарушений функционирования ограничивается социальным или профессиональным функционированием. Это не согласуется с формулировкой критерия нарушения функционирования для диагностики БДР и дистимии, который определяется нарушениями в социальной, профессиональной или *других важных областях функционирования*. Таким образом, состояние лиц, поддерживающих социальные контакты, но при этом не трудоустроенных и не выполняющих функции учащихся или родителей, может быть неправильно расценено как менее тяжелое, чем на самом деле.

Хотя умеренная степень тяжести и не определена конкретно, внутренняя логика формулировки описания средней степени тяжести имеет незначительный недостаток. Для легкой депрессии требуется небольшая степень выраженности симптомов, небольшой уровень дистресса и нарушений функционирования. И наоборот, для тяжелой депрессии требуется высокий уровень выраженности всех трех критериев. Таким образом, умеренная депрессия должна быть определена как находящаяся между легкими и тяжелыми уровнями выраженности симптомов, дистресса или нарушений функционирования (не *и/или*, как определяется в DSM-5).

Наконец, еще две переменные, которые считаются важными при обсуждении тяжести депрессии, – суицидальный риск и необходимость госпитализации – не рассматриваются при определении степени тяжести в DSM-5.

Какие доказательства подтверждают обоснованность подхода DSM-5 к определению степени тяжести подобным образом? Одно из реестровых популяционных близнецовых исследований, где участники перенесли большой депрессивный эпизод за год до интервью, показало, что три аспекта спецификатора тяжести – количество симптомов, тяжесть симптомов и степень нарушения функционирования – достоверно, хоть и незначительно, коррелировали⁴⁷. В связи с этим авторы пришли к выводу о многогранности и неоднородности определения тяжести по DSM.

Исследование амбулаторных пациентов с аффективными расстройствами⁴⁸, 84% из которых был выставлен диагноз большого депрессивного эпизода, показало, что количество симптомов БДР по DSM-IV имело слабую корреляцию с оценкой врачей по шкалам Clinical Global Impression (CGI)⁴⁹ и Global Assessment of Functioning (GAF)⁵⁰. Кроме того, оценка тяжести отдельных симптомов депрессии значительно коррелировала с баллами CGI и GAF, как и общее количество депрессивных симптомов. В небольшом исследовании стационарных пациентов с БДР было выявлено, что количество критериев БДР имело слабую корреляцию с Global Assessment Scale⁵¹. Kessler и соавт.⁵² проанализировали данные исследования National Comorbidity Study (NCS) и обнаружили, что, по сравнению с людьми, которые имели пять или шесть критериев БДР во время наиболее тяжелого эпизода депрессии, люди с семью-девятью критериями БДР имели более выраженные психосоциальные нарушения, больше эпизодов депрессии и большую длительность заболевания. Wakefield и Schmitz^{53,54} на основании изучения базы данных NCS и другого эпидемиологического исследования предположили, что в отношении прогноза о развитии большого депрессивного эпизода, необходимости профессиональной помощи, попытке самоубийства и госпитализации в психиатрическую больницу количество депрессивных симптомов было менее важным, чем их тип и особенности осложнений. Таким образом, количество симптомов не является адекватным показателем тяжести депрессии.

В МКБ-10⁵⁵ определяются три степени тяжести депрессии – легкая, умеренная и тяжелая – в зависимости от количества симптомов, тяжести симптомов, нарушений функционирования, уровня дистресса и, косвенно, типа симптомов. В отличие от DSM-5, в описании трех степеней тяжести отсутствует симметрия. Легкая депрессия определяется наличием двух или трех симптомов, которые нарушают функционирование, хотя пациент в таких случаях скорее всего продолжает большинство видов деятельности. Умеренная депрессия определяется четырьмя или более симптомами при наличии выраженных трудностей в повседневной деятельности. Тяжелая депрессия определяется «несколькими симптомами, которые выражены и вызывают дистресс, типично снижение самооценки и идеи бесполезности или вины. Суицидальные мысли и действия

являются частыми, и обычно присутствует ряд «соматических» симптомов».

Как и с определением спецификатора тяжести в DSM-5, спецификатор тяжести МКБ-10 мало исследовался, возможно, потому, что достоверность разграничения тяжести по степеням неудовлетворительна⁵⁶. Низкая достоверность неудивительна из-за неточности определений степеней тяжести⁵⁷.

Определения степеней тяжести официальных диагностических систем не использовались в исследованиях терапии. Точнее, практически во всех этих исследованиях степень тяжести определяется по результатам шкал – Шкалы Гамильтона для оценки депрессии (Hamilton Depression Rating Scale – HAMD)⁵⁸ или Шкалы депрессии Монтгомери–Асберга (Montgomery-Åsberg Depression Rating Scale – MADRS)⁵⁹. Таким образом, исследования терапии, как правило, не учитывают некоторые факторы, которые используются для характеристики тяжести, такие как уровень нарушения функционирования, уровень суицидального риска или подтип депрессии (т. е. наличие меланхолических черт или психотических симптомов)^{60,61}.

Шкалы для оценки тяжести депрессии

Тяжесть депрессии чаще всего определяется с помощью рейтинговых шкал, заполняемых в письменном виде, или шкал, оцениваемых клиницистами. Существуют различия оценочных инструментов по охватываемым временным рамкам (два наиболее распространенных временных периода – последние одна или две недели), по оценочным рекомендациям [большинство шкал используют оценку типа Лайкерт (Likert-type), основываясь на частоте симптомов, длительности или интенсивности] и по оцениваемым элементам.

Достаточно мало исследований было посвящено выявлению наиболее достоверных показателей тяжести депрессии. Как лучше определить тяжесть депрессии: числом симптомов (присутствующих или отсутствующих), частотой симптомов (например, каждый день, половину дней, несколько дней), персистированием симптомов (присутствуют все время, часто, иногда) или выраженностью симптомов (тяжелые, умеренные, легкие)? Для стандартизации оценки HAMD Williams и соавт.⁶² создали способ в виде таблицы для включения информации о частоте/персистенции и выраженности симптомов. Одно исследование, оценивающее важность учета как выраженности, так и частоты симптомов, выявило, что их интенсивность была лучшим предиктором тяжести, чем частота⁶³. При разработке шкалы для оценки депрессии Patient-Reported Outcomes Measurement Information System (PROMIS) Pilkonis и соавт.⁶⁴ рассматривали исследования, сравнивающие альтернативные варианты ответа, и заключили, что оценка частоты симптомов была более информативна, чем оценка их выраженности, однако в этих исследованиях оценивалась не депрессия. Таким образом, наиболее валидный формат шкал для оценки тяжести депрессии неизвестен и остается малоизученным.

Должно ли содержание шкалы тяжести основываться на диагностических критериях расстройства, включать другие симптомы депрессии, которые не являются компонентами диагностических критериев (например, низкая мотивация), или симптомы, которые часто встречаются у пациентов с депрессией, но являются определяющими признаками других расстройств (например, тревога, раздражительность)? И по какому стандарту следует судить, является ли один подход или шкала более достоверным показателем тяжести? Для построения шкал были использованы статистические подходы, такие как современная теория тестирования^{65,66}. Хотя инструменты, полученные на основе этого подхода, могут психометрически превосходить способы

оценки, основанные на диагностических критериях БДР, такие подходы не включают симптомы, которые уже давно считаются основными компонентами депрессии (нарушение аппетита и сна или риск суицида). Если оценка тяжести используется в клинических целях, а не только для изучения статистических данных, важно включать вегетативные симптомы, поскольку их наличие влияет на выбор терапии⁶⁷. Также следует оценивать риск суицида из соображений безопасности.

Несмотря на различия в устройстве шкал, их целевом назначении, охвате элементов и рекомендациях по оценке, общее заключается в том, что степень тяжести депрессии представлена суммой оценок отдельных элементов. Для всех шкал, кроме нескольких отдельных, все элементы оцениваются одинаково и вносят равный вклад в общую сумму баллов. Заметным исключением является шкала HAMD⁵⁸, которая включает в себя некоторые элементы с рейтингом от 0 до 2, а некоторые – с рейтингом от 0 до 4. Действительно, некоторые элементы отличаются по своему влиянию на разные домены депрессии⁶⁸. Некоторые оценочные инструменты были подвергнуты критике за мультидимENSIONALную структуру, поскольку одномерная конструкция тяжести депрессии способна лучше отразить эффективность лечения⁶⁹. Тем не менее все шкалы, даже мультидимENSIONALные с субшкалами, а также инструменты, которые первоначально предназначались для скрининга депрессии, а не оценки тяжести, имеют суммарную общую оценку, которая используется для обозначения тяжести расстройства.

Подход суммирования баллов основан на некоторых предположениях, которые не подтверждены эмпирически. Сложение баллов отдельных элементов для получения общей оценки в качестве показателя общей тяжести депрессии предполагает, что все симптомы являются одинаковыми индикаторами тяжести. Однако различные симптомы депрессии не одинаково коррелируют с глобальной оценкой тяжести клиницистами⁴⁸. С психометрической точки зрения, оценка отдельных элементов должна передавать достоверную информацию о всем спектре тяжести⁷⁰. Таким образом, пациенты с тяжелой депрессией должны чаще получать самые высокие баллы по симптомам, чем низкие или нулевые оценки, а пациенты с легкой депрессией должны чаще получать баллы, отражающие легкую степень тяжести, а не высокие показатели. Используя методы анализа данных современной теории тестирования, Santor и Coynе⁷⁰ продемонстрировали, что некоторые из элементов шкалы HAMD не соответствуют указанным положениям.

Действительно, шкалы, основанные на оценке частоты, редко соответствуют этим положениям, поэтому могут быть не лучшим методом оценки тяжести. Например, элементы опросника 9-item Patient Health Questionnaire (PHQ-9) оцениваются по 4-балльной шкале частоты симптомов за последние две недели (0 – не было совсем, 1 – несколько дней, 2 – больше половины дней, 3 – практически каждый день)⁷¹. Пациенты с БДР, вероятно, получают 3 балла по большинству присутствующих симптомов, потому что определение БДР подразумевает присутствие симптомов на протяжении минимум двух недель. Из-за эффекта измерительного «потолка» в условиях первичной медико-санитарной помощи пациент с БДР, который продолжает работать, получит ту же сумму баллов, что и пациент, госпитализированный из-за проблем с самообслуживанием. Хотя существуют исследования PHQ-9 с использованием подходов современной теории тестирования, в них рассматривались гетерогенные не чисто депрессивные психиатрические, медицинские или общие выборки^{72–78}. Нам неизвестны какие-либо исследования по оценке результатов PHQ-9 на выборке пациентов с депрессией для подбора терапии. Мы могли бы предположить, что в такой

выборке некоторые, а возможно, и многие элементы PHQ-9 будут сильно искажены из-за вышеупомянутого эффекта «потолка». Ни в одном исследовании не изучалось влияние различных подходов оценки на характеристики элементов шкалы оценки депрессии.

В подходе суммирования баллов подразумевается, что низкие оценки многих симптомов отражают степень тяжести, равную высоким оценкам меньшего числа симптомов. Например, тот, кто указывает, что в последние недели он редко испытывал такие симптомы, как подавленное настроение, бессонница, низкая самооценка, чувство вины, снижение концентрации внимания, утомляемость, психомоторное замедление, снижение аппетита, сложности в принятии решений и снижение интереса к повседневной деятельности, будет оценен той же степенью тяжести, что и тот, кто сообщит о ежедневно подавленном настроении, чувстве вины, чувстве собственной неполноценности и суицидальных мыслях, но будет отрицать все соматические и вегетативные симптомы депрессии. Аналогично, когда оценка элементов основана на выраженности симптомов, легкая выраженность многих симптомов предполагается равной значительной выраженности более ограниченного числа симптомов.

Подход суммирования баллов, при котором все оцененные элементы считаются равными, не основан на конкретной концепции тяжести. Если тяжесть болезни концептуализировать с точки зрения риска смертности, то следовало бы ожидать, что оценка суицидальных мыслей, чувства безнадежности и психомоторного возбуждения будет иметь больший вес, чем оценка нарушенной концентрации и усталости. С другой стороны, если тяжесть болезни концептуализировать с точки зрения нарушений функционирования, то можно ожидать, что элементы, оценивающие сниженную концентрацию и усталость, будут иметь большее значение, чем элементы, оценивающие снижение аппетита или чувство вины. Конечно, некоторые инструменты оценивают нарушения функционирования наряду с остальной симптоматикой^{63,71,79-81}, однако не было разработано инструментов, основанных на симптомах, путем изучения связи отдельных элементов с показателями нарушений функционирования и включения в шкалу только тех элементов, которые независимо связаны с данными нарушениями.

Несколько исследований были посвящены изучению связи между оценкой тяжести отдельных симптомов депрессии и множественными внешними показателями тяжести. Faravelli и соавт.⁴⁸ нашли четкие различия среди симптомов по их связи с оценкой по шкалам CGI и GAF. Кроме того, симптомы с наибольшей корреляцией с рейтингами CGI – такие как подавленное настроение, психическая заторможенность, нарушение концентрации внимания и ангедония – имели тенденцию к более значимым корреляциям с баллами GAF.

Большинство дискуссий о проблемах со шкалами по оценке депрессии сосредоточены на ограничениях применения их результатов^{69,82,83}. В то же время различные аспекты результирующих показателей могут представлять интерес и привести к развитию различных подходов в построении шкал. Некоторые способы оценки тяжести депрессии были созданы специально для отражения чувствительности к эффектам терапии^{59,84}. Ряд шкал связаны с критериями (симптомами), которые используются для диагностики депрессии^{71,79,85,86}, тогда как другие оценивают широкий спектр показателей, которые указываются пациентами как наиболее важные в оценке исхода⁸⁰, или оценивают ряд диагностических и связанных с ними симптомов депрессии⁸⁷. Описание построения шкалы обычно нацелено на содержание измерений и редко отражает причину выбора формата рейтинга. Например, при разработке шкалы Mul-

tidimensional Depression Assessment Scale Cheung and Power⁶⁸ было рассмотрено содержание пятнадцати шкал оценки депрессии и то, как разрабатываемая шкала будет устранять существующий пробел в содержании. Однако вопрос о форматах оценки и о том, почему для их шкалы был выбран формат оценки частоты симптомов, а не формат оценки выраженности симптомов, не обсуждался.

Одна из широко используемых клиницистами шкал оценки тяжести, MADRS, была разработана для точного отражения изменений состояния в исследованиях терапии⁵⁹. Элементы выбирались в том случае, если они были распространены у пациентов в начале лечения (т. е. имели распространенность более 70%) и показали наибольшее изменение от исходного уровня к 4-й неделе лечения, а изменение баллов от базового к 4-й неделе по симптому показало наибольшую корреляцию с изменением общей суммы баллов. Хотя в данном построении шкалы для указанной цели нет ничего изначально неправильного, это не должно быть основой для выбора элементов оценки тяжести депрессии, поскольку полученная шкала может быть смещена в сторону включения элементов, которые особенно чувствительны к изменениям для исследуемых препаратов. Конструкция MADRS была основана на реакции на мапротилин, amitриптилин и кломипрамин – препараты, которые в основном не применяются в настоящее время. Использование такого подхода для построения шкалы сейчас, когда назначаются другие препараты, может привести к созданию шкалы, которая лишь частично перекрывается с элементами, включенными в MADRS. Аналогичным образом, шкала HAM-D, которая была опубликована более 50 лет назад, подвергается критике за включение элементов, которые наиболее чувствительны к воздействию седативных препаратов, таких как трициклические антидепрессанты⁸⁸.

Таким образом, в то время как существует множество рейтинговых шкал для оценки депрессии и несколько исследований, изучающих их, остаются вопросы о том, как определить, является ли одна шкала более достоверным показателем тяжести депрессии, чем другая. Должна ли она быть основана на психометрическом анализе, отражающем одномерность? Будет ли «лучший» инструмент для оценки степени тяжести в большей степени коррелировать с индексами нарушений? Более значимо коррелировать с суицидальными мыслями? Будет ли он более прогностически значимым в отношении появления суицидального поведения в будущем? Более прогностически значимым в отношении будущей смертности в целом? Более прогностически значимым в отношении дальнейшего течения расстройства? Лучше дифференцировать пациентов с депрессией, которым требуется и не требуется госпитализация? Отражать больший размер эффекта в исследованиях терапии? Иметь свойство лучше разграничивать депрессию и тревогу, и, таким образом, являться «более чистым» показателем депрессии?

Проблема шкал для оценки депрессии: неопределенная валидность пороговых значений в определении степеней тяжести

Помимо вопросов о концептуализации тяжести и создании шкал, проблема тяжести депрессии заключается в несогласованности пороговых значений по шкалам симптомов, используемым для разграничения степеней тяжести, в особенности при тяжелой депрессии. Использование различных пороговых значений для определения групп тяжести затрудняет сравнение исследований по изучению влияния тяжести на выбор терапии.

DeRubeis и соавт.⁸⁹ провели мега-анализ четырех исследований, сравнивающих когнитивно-поведенческую терапию и медикаментозное лечение, и определили тяжелую депрессию как 20 или более баллов по шкале HAM-D,

состоящей из 17 пунктов. Аналогичным образом использовали значение в 20 баллов для определения тяжелой депрессии в недавнем мета-анализе плацебоконтролируемых испытаний флуоксетина и венлафаксина. Оба этих исследования ссылались на фундаментальную работу Elkin и соавт.⁹¹, чтобы обосновать свое определение тяжелой депрессии. Однако Elkin и соавт. не приводят эмпирических доказательств в пользу этих значений и, по сути, не сообщают о пациентах, набравших более 20 баллов по шкале HAMD, в абсолютном смысле (т. е. с тяжелой депрессией), а упоминают их лишь в относительном (т. е. с более тяжелой депрессией, чем пациенты, набравшие 20 баллов и менее).

В мета-анализе Kirsch и соавт.⁹², посвященном изучению влияния тяжести на различия в терапии «антидепрессант – плацебо», авторы отметили, что средние баллы по шкале HAMD в исследованиях эффективности антидепрессантов были в спектре очень тяжелых (т. е. более 23 баллов), основываясь на Handbook of Psychiatric Measures⁹³ Американской психиатрической ассоциации (АПА), для всех, кроме двух из 35 включенных в анализ исследований. В предыдущем анализе исследований эффективности антидепрессантов по базе данных Управления по санитарному надзору за качеством пищевых продуктов и медикаментов (Food and Drug Administration – FDA) США Khan и соавт.⁹⁴ разделили исследования на три группы на основе оценок по шкале HAMD до начала лечения (<24; 25–27; >28) без обоснования указанных пороговых значений. Fournier и соавт.⁹⁵ использовали пороговые значения, рекомендованные в Handbook of Psychiatric Measures⁹³ АПА для определения степеней тяжести по шкале HAMD (от легкой до умеренной – менее 18; тяжелая – 19–22; очень тяжелая – более 23). В отличие от этих исследований и рекомендаций АПА, большинство фармакотерапевтических исследований использовали значение в 25 баллов по шкале HAMD из 17 пунктов для определения депрессии тяжелой степени^{96–101}, и указанное пороговое значение было рекомендовано некоторыми экспертами^{102–104}. Таким образом, депрессия тяжелой степени не имеет однозначного определения.

Фундаментальным положением для исследований значенности степеней тяжести в лечении является обоснованность пороговых значений для определения категорий тяжести по шкале HAMD. Ни в одном из разделов обсуждения мета-анализов и объединенных анализов сообщений о степени тяжести и результатах лечения не поднимались вопросы о пороговых значениях, используемых для определения степени тяжести. В Handbook of Psychiatric Rating Scales⁹³ АПА процитировано только два небольших исследования в поддержку пороговых значений для определения степеней тяжести, и ни одно из этих исследований не проводилось для создания рекомендаций АПА. Одно исследование изучало валидность соответствующих суммарных баллов по шкале HAMD для Schedule for Affective Disorders and Schizophrenia¹⁰⁵. В этом исследовании не пытались определить пороговые баллы по HAMD, указывающие на степень тяжести. Во втором исследовании изучалась связь между баллами HAMD и глобальными рейтингами тяжести у 59 стационарных пациентов с депрессией¹⁰⁶. Авторы не выводили (и не рекомендовали) пороговые значения, соответствующие степеням тяжести. Таким образом, неясно, почему значение в 19 баллов было рекомендовано в руководстве АПА для выявления тяжелой депрессии. UK National Institute for Health and Clinical Excellence (NICE) рекомендовал значение в 23 балла для определения тяжелой депрессии по шкале HAMD, хотя эта рекомендация не была подкреплена ссылками на исследования¹⁰⁷.

Из-за ограниченного количества эмпирических исследований, устанавливающих пороговые баллы для степеней тяжести по шкале HAMD, и значения, придаваемого степени тяжести руководствами по лечению, наша клиническая

исследовательская группа также рассмотрела эту проблему среди 627 амбулаторных пациентов с БДР, которые были оценены по шкале CGI¹⁰⁸. Пороговое значение по шкале HAMD, которое максимизировало сумму чувствительности и специфичности, составило 17 баллов относительно легкой и умеренной депрессии и 24 балла относительно умеренной и тяжелой депрессии. На основе обзора имеющихся фактических данных, а также рекомендаций о том, что для определения ремиссии следует использовать значение в 7 баллов, мы рекомендовали следующие диапазоны для оценки тяжести по шкале HAMD из 17 пунктов: отсутствие депрессии (0–7), легкая депрессия (8–16), умеренная депрессия (17–23) и тяжелая депрессия (>24).

Каждое из вышеперечисленных исследований основывалось на суждениях клиницистов о степени тяжести. Ограничением этих исследований является то, что неизвестно, на основании чего были сделаны выводы о тяжести. Считались ли некоторые симптомы депрессии лучшими показателями тяжести, чем другие симптомы? Например, придается ли большее значение симптомам, характерным для меланхолической или эндогенной депрессии, в оценке по CGI? Не слишком ли непропорционально влияет оценка риска суицида на рейтинги клиницистов? Учитывают ли врачи психосоциальные нарушения при составлении рейтинга CGI? Нам не известно о каких-либо исследованиях, которые пытались бы получить диапазоны тяжести по шкале HAMD или любой другой шкале для оценки депрессии по этому вопросу, основанные на степени нарушений или риске суицида.

Дополнительная проблема со шкалами для оценки симптомов депрессии: различные шкалы неодинаково распределяют пациентов по степени тяжести

В клинической практике опросники с самооценкой предпочтительнее шкал, оцениваемых врачом, поскольку работа с ними требует меньше времени. Если для классификации пациентов по категориям тяжести необходимо использовать шкалы самооценки и если рекомендации по лечению должны частично основываться на классификации тяжести, то важно, чтобы разные шкалы классифицировали людей одинаково. Однако, поскольку содержание инструментов для оценки различается, неудивительно, что между ними существуют значительные различия.

Cameron и соавт.¹⁰⁹ сравнили классификацию тяжести по PHQ-9 и Госпитальной шкале тревоги и депрессии (Hospital Anxiety and Depression Scale – HADS) в выборке пациентов из звена первичной медико-санитарной помощи, направленных врачами общей практики Великобритании к специалисту в области психического здоровья¹¹⁰. Никакой информации о психиатрических диагнозах пациентов предоставлено не было. В исследовании обнаружили, что по результатам PHQ-9, в сравнении с HADS, вдвое большее число пациентов было включено в спектр тяжелой степени. Другие исследования, сравнивающие PHQ-9 и HADS на выборках пациентов, выявили схожие результаты^{111,112}. Однако эти исследования нуждаются во внешней валидации, и поэтому неясно, превышает ли PHQ-9 степень тяжести, или это HADS ее занижает. Во втором исследовании Cameron и соавт.¹⁰⁷ включили второе издание Beck Depression Inventory (BDI-II)¹¹³ вместе с PHQ-9 и HADS, и также оценивали пациентов с помощью HAMD. Участниками были пациенты первичной медицинской практики, которым был поставлен диагноз депрессии их терапевтом. PHQ-9 и BDI-II завышали степень тяжести по сравнению с HAMD, а шкала HADS ее занижала.

Нам известно только об одном исследовании, в котором сравнивались шкалы самооценки в выборке амбулаторных пациентов с БДР¹¹⁴. Наша клиническая исследовательская

группа сравнила классификацию тяжести по трем шкалам, которые оценивают диагностические критерии симптомов DSM-IV/DSM-5 для БДР: Clinically Useful Depression Outcome Scale (CUDOS)⁷⁹, Quick Inventory of Depressive Symptomatology (QIDS)⁸⁵ и PHQ-9⁷¹. Пациенты также были оценены по шкале HAM-D из 17 пунктов. В исследовании амбулаторных пациентов с депрессией мы обнаружили, что корреляции между баллами шкалы HAM-D и всеми тремя шкалами самооценки были почти идентичны, а средняя корреляция между тремя шкалами самооценки составляла 0,73. Однако шкалы достоверно различались в распределении пациентов по категориям тяжести. Примерно треть пациентов были отнесены в диапазон легкой степени тяжести по шкалам HAM-D и CUDOS, в то время как по шкалам PHQ-9 и QIDS приблизительно 10% пациентов имели депрессию легкой степени. По шкалам CUDOS и HAM-D депрессия средней тяжести была самой многочисленной категорией, тогда как по шкалам PHQ-9 и QIDS большинство пациентов были отнесены в группу депрессии тяжелой степени. Большинство пациентов в диапазоне средней степени тяжести по шкале HAM-D находились в диапазоне тяжелой степени по шкалам PHQ-9 и QIDS. Значительно меньшее число пациентов были классифицированы как тяжело депрессивные по шкале CUDOS, по сравнению с PHQ-9 и QIDS.

С учетом того что три шкалы самооценки значительно коррелируют друг с другом и одинаково коррелируют с HAM-D, что же тогда может объяснить заметные различия между шкалами аналогичного содержания в распределении пациентов по группам тяжести?

Пороговые значения трех шкал для определения групп тяжести были получены разными способами, и это, вероятно, объясняло различия между шкалами в классификации тяжести. Например, Kroenke и соавт.⁷¹ указали, что пороговые значения по шкале PHQ-9 были выбраны по прагматической причине – чтобы облегчить их запоминание врачами. Они также отметили, что альтернативные пороговые значения не улучшают связь между повышением тяжести по PHQ-9 и показателями валидности конструкции. При выборе пороговых значений для определения диапазонов тяжести по PHQ-9 разработчики данного опросника не учитывали потенциальное влияние широты каждой из групп тяжести, с помощью которой были определены диапазоны тяжести, и то, как это может повлиять на рекомендации по лечению, содержащиеся в официальных руководствах.

Kroenke и соавт.⁷¹ указали, что, когда группы тяжести, основанные на разных пороговых значениях, одинаково коррелируют с внешними переменными, пороговые значения могут быть выбраны на основе простоты их запоминания. Мы не согласны с этой аргументацией. Для всех шкал, измеряющих тяжесть депрессивных симптомов, пороговые значения, отличающие пациентов с легкой, средней и тяжелой депрессией, не представляют четко разграниченных критериев, разделяющих подтипы тяжести. Как и в других областях психопатологии, выраженность депрессии лучше соответствует размерной, чем категориальной модели классификации¹¹⁵. Таким образом, альтернативные пороговые значения для классификации групп тяжести, вероятно, будут валидны при сравнении групп по внешней переменной, такой как психосоциальное функционирование. Однако следует с осторожностью подходить к выбору пороговых значений, поскольку они влияют на относительный диапазон каждой из категорий тяжести.

Если клиницисты должны следовать рекомендациям официальных руководств по лечению и основываться в первоначальном выборе терапии на тяжести депрессии, то необходимо иметь унифицированный метод определения тяжести депрессии. Отмеченные различия между стандар-

тизированными шкалами самооценки при классификации амбулаторных пациентов с депрессией на группы тяжести свидетельствуют о наличии проблемы с использованием таких инструментов для классификации тяжести депрессии. Если следовать официальным рекомендациям по лечению, то использование завышающих тяжесть инструментов, таких как QIDS и PHQ-9, приведет к большему применению фармакотерапии относительно психотерапии в качестве лечения первой линии для БДР. Таким образом, следует проявлять осторожность при использовании этих шкал для определения выбора лечения до тех пор, пока пороговые значения для определения диапазонов тяжести не будут лучше определены эмпирическим путем.

Значение тяжести депрессии для терапии: рекомендации официальных руководств

Несмотря на вышеупомянутые проблемы с концептуализацией тяжести депрессии и определением границ степеней тяжести по шкалам, тяжесть депрессии является важным фактором при принятии решений о лечении. Тяжесть депрессии влияет на рекомендации по лечению в официальных руководствах. В третьем издании рекомендаций APA по лечению БДР рекомендуются как психотерапия, так и фармакотерапия в качестве монотерапии при депрессии легкой и средней степени тяжести и фармакотерапия (с психотерапией или без нее) при тяжелой депрессии¹. Согласно обновленному руководству NICE по лечению и купированию депрессии, использование антидепрессантов в качестве первоначального варианта лечения легкой депрессии не рекомендуется, а для умеренной и тяжелой депрессии рекомендуется медикаментозное лечение вместе с эмпирически обоснованной психотерапией². Как сообщили van der Lem и соавт.¹¹⁶, голландское руководство по лечению также рекомендует фармакотерапию в качестве терапии первой линии для пациентов с тяжелой депрессией и фармакотерапию или психотерапию для пациентов с депрессией легкой и средней степени тяжести. Хотя рекомендации, содержащиеся в этих руководствах, не являются полностью унифицированными, они единогласно рекомендуют фармакотерапию в качестве метода выбора терапии тяжелой депрессии.

Важность тяжести расстройства для терапии была изучена несколькими различными способами. Существуют контролируемые исследования, исследования эффективности, объединенные анализы и мета-анализы, посвященные изучению влияния тяжести расстройства на тип лечения^{117–122}, сравнению лечения по диапазонам тяжести^{99,123–127}, сравнению препаратов и плацебо по диапазонам тяжести^{128,129}, сравнению психотерапии с группами контроля по диапазонам тяжести^{130,131}, сравнению лечения среди пациентов с тяжелой депрессией^{96,101,102,132} и оценке того, прогнозирует ли тяжесть краткосрочные результаты^{42,133–135}, резистентность к лечению¹³⁶, долгосрочные результаты^{40,137–139} и рецидивы³⁸.

Тяжесть депрессии и фармакотерапия

В последнее десятилетие поднимаются вопросы о том, эффективны ли селективные ингибиторы обратного захвата серотонина (СИОЗС) и другие антидепрессанты нового поколения при легкой депрессии. Khan и соавт.⁹⁴ проанализировали 45 клинических исследований по базе данных FDA и обнаружили, что в исследованиях со средним исходным баллом ≤ 24 по шкале HAM-D из 17 пунктов мало доказательности большей эффективности антидепрессантов относительно плацебо. При этом в исследованиях со средним исходным баллом ≥ 28 по шкале HAM-D имелись явные доказательства того, что препараты превосходили по эффективности плацебо. Kirsch и соавт.⁹² аналогичным образом проанализировали базу данных FDA, а также

изучили эффективность антидепрессантов в зависимости от среднего исходного балла по шкале HAMD. Их результаты в основном воспроизвели результаты Khan и соавт.⁹⁴ о том, что разница между препаратом и плацебо была наиболее выражена в исследованиях с наиболее высоким исходным показателем тяжести (т. е. >28 баллов по HAMD). Kirsch и соавт.⁹² выявили, что антидепрессанты были достоверно более эффективны, чем плацебо, в когортах с менее тяжелой депрессией, но они расценили эту разницу как незначительную и клинически несущественную.

В отличие от анализа базы данных FDA, проведенного Kirsch и соавт.⁹² и Khan и соавт.⁹⁴, Fournier и соавт.⁹⁵ объединили данные об отдельных пациентах из шести опубликованных исследований. Kirsch и соавт. и Khan и соавт. использовали агрегированные средние баллы для всего исследования в качестве единицы анализа, т. е. они сравнивали исследования с различными средними исходными баллами тяжести. Проблема такого подхода заключается в том, что группа пациентов со средним баллом в диапазоне тяжелой степени также будет включать некоторых пациентов с баллами в диапазоне легкой и средней тяжести. Аналогичным образом, группа пациентов со средним баллом в диапазоне легкой или средней тяжести будет включать в себя некоторых пациентов с баллами в диапазоне тяжелой степени. Объединение данных отдельных пациентов позволяет избежать проблемы неправильного распределения групп тяжести на уровне отдельных пациентов. Fournier и соавт.⁹⁵ также сделали вывод о том, что различия между препаратом и плацебо были клинически значимыми только для пациентов с тяжелой депрессией, и выявили небольшую величину эффекта для пациентов с легкой и умеренной депрессией.

Совсем недавно были проведены другие исследования по анализу объединенных данных на уровне пациента (а не агрегированные данные из исследования). Использовались базы данных фармацевтических компаний, в анализ включались все исследования продукта, тем самым избегая предвзятости, присущей изучению только опубликованных исследований¹⁴⁰. Согласно результатам трех крупных объединенных анализов опубликованных и неопубликованных исследований, в которые было включено от 4000 до 10 000 пациентов, антидепрессанты эффективны во всем диапазоне тяжести^{90,129,141}. Эти анализы и полемика относительно эффективности антидепрессантов подчеркивают необходимость оценки тяжести в клинической практике.

Тяжесть депрессии и фармако- или психотерапия в качестве лечения первой линии

Второй важный вопрос лечения, связанный с тяжестью расстройства, заключается в том, следует ли рассматривать тяжесть депрессии в качестве основания для рекомендации фармако- или психотерапии в качестве лечения первой линии. Если быть точнее, вопрос заключается в том, должны ли пациенты с тяжелой депрессией получать преимущественно фармакотерапию. Другой связанный с этим вопрос: является ли психотерапия полезной для пациентов с тяжелой депрессией?

Тяжесть симптомов в качестве предиктора ответа на лечение является предметом продолжающихся дебатов с момента публикации результатов US National Institute of Mental Health Treatment of Depression Collaborative Research Program (TDCRP), предполагающих, что психотерапия не столь эффективна, как препараты в качестве купирующей терапии при тяжелой депрессии. Первый мета-анализ исследований, непосредственно сравнивающих психотерапию и фармакологические вмешательства, охватывал 30 опубликованных исследований, включавших более 3000 пациентов¹⁴³. Мета-регрессионный анализ, изучающий связь размера эффекта со средними исходными балла-

ми по шкалам HAMD или BDI, не выявил доказательств того, что исходный уровень тяжести связан с различными результатами лечения. Сравнение размеров эффекта в исследованиях с исходными баллами по шкале HAMD ниже 20 и выше 20 также не выявило различий.

Мета-анализ 132 контролируемых исследований психотерапии на выборке из более чем 10 000 пациентов выявил, что большая исходная средняя тяжесть симптомов не была предиктором более плохого ответа¹³⁰. Совсем недавно Weitz и соавт.¹⁴⁴ объединили индивидуальные данные пациентов из 16 исследований, сравнивающих антидепрессанты и когнитивно-поведенческую терапию. Они определили тяжелую группу в соответствии с рекомендациями APA (HAMD ≥ 19) и NICE (HAMD >23). Большая тяжесть была связана с достоверно более низкими показателями ремиссии (но не показателями ответа на терапию) в условиях как фармакотерапии, так и психотерапии. Тяжесть расстройства не была связана с различными результатами лечения, что подтвердило результаты предварительного объединенного анализа на основе меньшего числа исследований⁹⁹. В контрольном исследовании авторы провели объединенный анализ, сфокусированный на пяти исследованиях, в которых использовалось плацебо в качестве контроля¹³¹. Результаты были сопоставимы с большим объединенным анализом: исходная тяжесть симптомов не была связана с изменением тяжести симптомов от исходного к конечному уровню между группами с когнитивно-поведенческой терапией и группами, принимавшими плацебо.

Таким образом, результаты более поздних мета-анализов, основанных на распределении по группам тяжести в соответствии со шкалами оценки симптомов, не согласуются с официальными терапевтическими руководствами, которые рекомендуют фармакотерапию в качестве лечения первой линии при тяжелой депрессии.

ТЯЖЕСТЬ РАССТРОЙСТВ ЛИЧНОСТИ

Степень тяжести является важной характеристикой и для РЛ, хотя нынешние диагностические системы не содержат формальных способов оценки их тяжести. Пациенты с РЛ, признанные «тяжелыми», с большей вероятностью имеют другие коморбидные психиатрические диагнозы, в особенности аффективных, тревожных расстройств, зависимости от психоактивных веществ¹⁴⁵ и других РЛ¹⁴⁶. Так называемые «тяжелые пациенты» часто получают лечение в течение долгого времени¹⁴⁷⁻¹⁴⁹, демонстрируют более высокие показатели частоты госпитализации и суицидальных попыток¹⁵⁰, а также более часто наносят самоповреждения¹⁵¹. Они с большей вероятностью оказываются арестованы, чаще не в состоянии удерживаться на работе и чаще терпят неудачу в отношениях¹⁵². Общеизвестно, что они могут представлять опасность для окружающих, поэтому следует выявлять их как можно раньше для своевременного лечения^{3,4,153}.

Тем не менее вопрос, что означает «тяжелое» РЛ, остается открытым. Тяжесть оценивалась путем подсчета общего числа диагнозов коморбидных РЛ, при этом более высокая коморбидность указывала на большую тяжесть^{152,154-156}. Однако это, вероятно, в большей степени отражает тяжесть патологии личности в целом, а не тяжесть конкретного РЛ. В более тяжелых случаях патология личности может определяться сложностью конкретного случая и спецификой сопутствующих заболеваний. Основной раздел DSM-5 (т. е. Раздел II) определяет РЛ в рамках одного из трех кластеров. Tuger и Johnson¹⁵⁷ предположили, что люди с коморбидными РЛ из разных кластеров имеют большую тяжесть расстройства, чем люди с коморбидными РЛ из одного кластера. Авторы идентифицируют диссоциальное РЛ как наиболее тяжелое на основе риска для окружающих. Таким образом, в наиболее тяжелых случаях у паци-

ентов будет диагностировано диссоциальное РЛ, а также РЛ из других кластеров. При использовании этой модели тяжесть РЛ была ассоциирована с поведенческими расстройствами, преступным поведением, бродяжничеством, помещением в места лишения свободы, статусом безработного и делинквентным поведением в детстве.

Тяжесть конкретного РЛ может быть измерена путем подсчета числа критериев, удовлетворяющих диагнозу. Например, в случае пограничного РЛ пациенты, удовлетворяющие девяти критериям, будут расценены как более тяжелые, чем те, которые удовлетворяют только пяти критериям¹⁴⁷. Однако результаты нашей клинической исследовательской группы не смогли подтвердить эту гипотезу, не обнаружив различий в сопутствующей патологии или психосоциальном функционировании на основе подсчета количества критериев для пациентов с диагнозом пограничного РЛ¹⁵⁸. Кроме того, тяжесть может быть определена частотой симптомов. Например, пациенты с пограничным РЛ, которые наносят самоповреждения несколько раз в день, будут оценены как более тяжелые, чем те, которые делают это раз в месяц¹⁵¹.

Определенные РЛ даже были обозначены как более или менее тяжелые, чем другие. Kernberg и Caligot¹⁵⁹ ранжировали РЛ от «более тяжелых» (таких как пограничное РЛ) до менее тяжелых (например, ананкастное РЛ, зависимое РЛ). Также предлагалось концептуализировать РЛ с использованием совокупности патологических черт личности. С этой точки зрения, «тяжелый» симптом или признак РЛ может быть определен как статистически крайне выраженный или существующий только у очень небольшой части населения¹⁶⁰.

Исследования лечения «тяжелых» расстройств личности в первую очередь акцентируют внимание на характеристиках симптомов (частота, персистенция, выраженность) и нарушениях функционирования (социальное/профессиональное) или исходах, таких как тюремное заключение^{161–163}. Maden и Tugel¹⁶² выявляют категорию «опасных и тяжелых» РЛ, которая характеризуется высоким риском причинения непоправимого вреда другим. Как это ни странно, первый критерий наличия «опасного и тяжелого» РЛ диагностируется с помощью понятия «тяжелое расстройство личности», которое само по себе остается неопределенным. Авторы не уточняют, что означает тяжесть на уровне критериев, хотя, как представляется, это определение является правомочным и относится в первую очередь к психопатии, а не к РЛ в традиционном понимании.

Тяжесть расстройств личности и функционирование

Хотя в литературе по РЛ тяжесть определялась различными способами, появляется общий консенсус о том, что тяжесть РЛ неразрывно связана с уровнем дезадаптации^{164–169}. Многие признают, что крайне выраженных признаков или симптомов недостаточно для диагностики РЛ или для определения степени тяжести. Скорее суть заключается в наличии крайне выраженных черт личности при наличии нарушений адаптации, связанных с этими чертами. В отличие от соматических заболеваний или даже депрессии, которые больше основаны на наличии симптомов, диагноз РЛ тесно переплетается с адаптацией. Как и депрессия, РЛ по определению должны приводить к «дистрессу или дезадаптации», которые нужно диагностировать³³. Однако, в отличие от депрессии, диагностические симптомы РЛ включают как аффективный/когнитивный/эмоциональный, так и функциональный компоненты. Например, снижение профессиональной и финансовой продуктивности включено в критерии диссоциального РЛ, а нарушения в социальной сфере и неспособность уделять время досугу являются частью критериев ананкастного РЛ.

Взаимосвязь между нарушениями функционирования и личностью приводит многих к выводу о том, что тяжесть РЛ представляет собой сочетание крайне выраженных черт личности и дезадаптации, связанной с ними^{165,169}. В действительности наличие дезадаптации настолько фундаментально для диагностики РЛ и его тяжести, что некоторые авторы утверждают: оценка крайне выраженных признаков или симптомов неоправдана^{170–173}. Таким образом, не обязательно наличие крайне выраженных симптомов, если имеется достаточная дезадаптация. Однако дезадаптация обусловлена наличием личностных черт, даже если они выражены не в крайней степени. Например, используя многоосевую структуру DSM-IV, Livesley¹⁷⁴ предложил определять наличие РЛ по Оси I и кодировать черты личности отдельно по Оси II. Widiger и Trull¹⁶⁹ предложили аналогичную модель, но с использованием результирующего балла по шкале GAF на Оси V в качестве оценки степени тяжести.

Вместе эти модели сходятся на определении тяжести как общего, адаптивного нарушения внутриспсихической системы, необходимой для выполнения повседневных жизненных задач¹⁶⁶. Хотя конкретные сферы нарушений различаются, существуют общие нарушения в трех широких областях: формирование идентичности, самоконтроль (или направленность деятельности) и межличностные отношения¹⁶⁴. Тем не менее некоторые исследования показывают, что патологические черты личности и функционирование настолько тесно переплетены, что не могут представлять собой отдельные домены¹⁷⁵.

Тяжесть расстройств личности по DSM-5 и МКБ-10

В основном Разделе II DSM-5 нет четкого упоминания о степени тяжести РЛ. Однако общее описание РЛ включает показатели тяжести, общие для других расстройств. Например, отмечается, что РЛ характеризуются негибкостью, неадаптивностью, тотальностью и ассоциированы с «клинически значимыми» дезадаптацией или субъективным дистрессом. Дезадаптация является показателем тяжести многих соматических и психических расстройств; тотальность – показателем тяжести депрессии; а субъективный дистресс указывает на тяжесть аффективных расстройств и сексуальной дисфункции. В настоящее время не существует официального метода определения тяжести РЛ в DSM-5.

Раздел III (Emerging Measures and Models) DSM-5 включает альтернативную модель диагностики РЛ. Диагноз определяется сочетанием выраженности дисфункции и заостренных личностных черт, а тяжесть определяется главным образом дисфункцией, связанной с «заостренными» чертами³³. Эта модель не определяет способ оценки тяжести в целом, но для диагностики требуется «умеренное или большее нарушение функционирования». Нарушение функционирования операционализируется как один из пяти уровней, крайний из которых свидетельствует о тяжелой патологии личности. Для оценки нарушений функционирования и, следовательно, тяжести РЛ предлагается шкала Level of Personality Functioning Scale (LPFS). Оценка составляется для внутрличностного (идентичность и саморегуляция) и межличностного (эмпатия и близость) функционирования. Уровни включают в себя: 0 (незначительное нарушение или отсутствие нарушения), 1 (некоторые нарушения), 2 (умеренные нарушения), 3 (тяжелые нарушения), 4 (крайне выраженные нарушения). Лица с крайне выраженными нарушениями описываются как имеющие неопределенную, размытую идентичность и саморегуляцию с неадаптивной «Я-концепцией» и полностью лишенные способности к межличностным взаимодействиям.

Интересно, что Раздел III DSM-5 также включает в себя обсуждение дополнительного способа оценки личностных

черт, Personality Inventory for DSM-5¹⁷⁶. Его элементы связаны с личностными чертами, однако он определяет общий суммированный балл, измеряющий «общую дисфункцию личности». Идентификация крайне выраженных признаков в качестве определяющих дисфункцию необычна, но не противоречит значительному перекрытию между функционированием и признаками/симптомами РЛ, о котором сообщается в литературе¹⁷⁵. Тем не менее это говорит о том, что крайне выраженные признаки указывают на крайнюю степень дисфункции, которая является основным показателем тяжести в этой модели.

Как и в DSM-5, в МКБ-10 нет упоминаний о тяжести РЛ. Вместе с тем было опубликовано несколько документов по существенным изменениям, предлагаемым для МКБ-11. В частности, первичной будет классификация РЛ, основанная на тяжести нарушений. Описание признаков или черт РЛ возможно, но не требуется для диагностики^{3,4}.

В соответствии с большим количеством литературных источников, предложенные в МКБ-11 изменения концептуализируют тяжесть главным образом как дисфункцию или проблемы, связанные с личностью индивида. Предлагаются пять степеней тяжести, хотя они немного отличаются от аналогичных в DSM. В целом, степень тяжести определяется, во-первых, распространенностью нарушений (повсеместно или ограниченные ситуациями), а во-вторых, количеством «проблемных» черт личности (множественные или единичные). При самой высокой степени тяжести также оценивается риск причинения вреда себе или окружающим. Таким образом, наиболее тяжелые случаи прежде всего определяются нарушениями в функционировании. Симптомы/черты и риск причинения вреда вторичны, но также учитываются. В отличие от альтернативной модели DSM-5, дисфункция «самости» (self) и идентичности не включена в оценку тяжести^{3,4}. На момент подготовки данной статьи МКБ-11 еще не была опубликована, поэтому определения следует считать предварительными. Тем не менее акцент на дисфункции посредством оценки тяжести критиковался за недостаточный объем исследований, которые бы установили надежность и валидность такого метода¹⁷⁷.

Оценка тяжести расстройств личности

Еще в 1996 г. Tuger и Johnson¹⁵⁷ разработали пятибалльную шкалу оценки тяжести расстройств, похожую на предложенную в МКБ-11. Оценка была основана на информации, полученной инструментом для оценки личностных черт, Personality Assessment Schedule (PAS)¹⁵³. Таким образом, тяжесть больше определялась выраженностью черт, чем функционированием. PAS также используется для классификации индивидов на четыре категории, предложенные Tuger и Johnson¹⁵⁷: отсутствие РЛ, «сложная личность», простое РЛ, сложное РЛ. Обозначения степени тяжести по PAS преимущественно основаны на встречаемости категорий DSM-IV и ICD-10 и использовались в исследованиях, прогнозирующих исход лечения, хотя и со смешанными результатами¹⁷⁸. Инструмент General Assessment of Personality Disorder¹⁷⁹ использовался в качестве индекса тяжести в многочисленных исследованиях и предлагает две основные шкалы тяжести – патология «самости» (self) и межличностные проблемы – обе из которых отражают нарушения функционирования, определенные DSM-5^{164,180,181}. Аналогичным образом, Severity Indices of Personality Problems¹⁷³ определяет тяжесть как сочетание нарушенной «самости» (self) и межличностного взаимодействия.

Для отдельных РЛ существует относительно немного инструментов для оценки тяжести, и они в основном сосредоточены на пограничном РЛ. Например, Borderline Personality Disorder Severity Index (BPDSI)^{151,182} является структурированным клиническим интервью, которое опе-

рационализирует тяжесть в первую очередь частотой поведенческих симптомов пограничного РЛ в течение предыдущих трех месяцев. Частота симптомов оценивается от 0 (никогда) до 10 (ежедневно). Степень тяжести усредняется по результатам, давая оценку тяжести как для отдельных критериев пограничного РЛ, так и для расстройства личности в целом. Таким образом, BPDSI в значительной степени оценивает тяжесть как функцию от частоты симптомов, хотя многие из его элементов содержат вопросы о поведении, которое подразумевает дезадаптивные последствия (например, отсутствие работы).

Учитывая, что тяжесть патологии личности часто связана с нарушениями функционирования, исследования исходного лечения РЛ часто ориентированы на степень, в которой различные подходы к лечению (например, диалектическая поведенческая терапия, терапия, основанная на ментализации, сфокусированная на переносе психотерапии) улучшают повседневное функционирование и предотвращают дезадаптивное поведение^{147,183,184}. Например, в обширной литературе по лечению пограничного РЛ изменения часто оцениваются с помощью таких инструментов, как Zanarini Rating Scale for Borderline Personality Disorder¹⁸⁵ и Barratt Impulsiveness Scale¹⁸⁶. Тем не менее сокращение количества суицидальных попыток, уменьшение аутоагрессии и зависимости от неотложной службы психиатрической помощи часто являются основными показателями исхода лечения, так же как улучшения в поддержании значимых межличностных связей и улучшения, связанные с работой^{147,183,184,187,188}.

Хотя в литературе, посвященной лечению РЛ, основное внимание уделяется лечению пограничного РЛ, другие РЛ также исследуются, при этом нарушения функционирования были определены как показатели результатов лечения. Например, сфокусированная на переносе психотерапия оказалась несколько более эффективной для пациентов с коморбидными нарциссическим и пограничным РЛ. Этот подход к лечению подчеркивает межличностное функционирование в личных и трудовых отношениях при оценке результатов¹⁸⁹. Исследования лечения диссоциального РЛ ориентированы на последующее употребление психоактивных веществ и возможные аресты¹⁹⁰. Таким образом, среди различных РЛ результаты лечения и снижение «тяжести» определяются не только редукцией симптомов, но и уменьшением специфического опасного поведения (например, аутоагрессии) и развитием межличностных взаимоотношений и специфического просоциального поведения (например, сохранение работы).

ТРАНСДИАГНОСТИЧЕСКИЕ МОДЕЛИ И ТЯЖЕСТЬ: СПЕКТРЫ ПСИХОПАТОЛОГИИ

Многие из указанных выше вопросов о том, например, как сравнивать достоверность шкал депрессии при измерении степени тяжести, также применимы к определению того, предполагают ли различные диагнозы дифференциальные уровни тяжести. Вероятность удовлетворения критериев разных диагнозов, по всей видимости, определяется исходными генетическими механизмами^{191,192}. Это важно, учитывая, что лица, отвечающие критериям одного диагноза, скорее всего, отвечают критериям нескольких других¹⁹³, таким образом, что различные диагнозы могут быть проявлениями исходных спектров (например, диссоциальное РЛ, нарциссическое РЛ и наркомания вместе отражают спектр экстернализации).

Исследования, посвященные изучению связей между различными интернализационными диагнозами, характеризующимися субъективным дистрессом и страхом, показывают, что такие пациенты могут «лучше» соответствовать критериям такого диагноза, как БДР, по сравнению с более «тяжелыми» расстройствами, такими как генерализован-

ное тревожное и паническое расстройства¹⁹⁴. Иными словами, соответствие критериям генерализованного тревожного или панического расстройства отражает более высокое положение на дименсии интернализации, чем просто соответствие критериям БДР. Интересно, что Krueger и Finger¹⁹⁴ также обнаружили, что высокая позиция по дименсии интернализации была значительно связана с пожизненным количеством стационарных госпитализаций и психосоциальным функционированием в прошлом месяце.

В других недавних исследованиях «тяжесть» на спектре интернализации также связана с ключевыми результатами. Например, Eaton и соавт.¹⁹⁵ обнаружили, что вероятность соответствия критериям для различных депрессивных расстройств, тревожных расстройств и биполярного расстройства может быть представлена исходным континуумом. Лица с высокими баллами по этой дименсии, которые будут характеризоваться как имеющие более «тяжелые» уровни психопатологии интернализации, скорее всего, соответствуют критериям многих диагнозов и будут сообщать о широком спектре симптомов (например, сниженное настроение, беспокойство, трудности концентрации, раздражительность), характеризующих различные диагнозы этой дименсии по DSM.

Eaton и соавт.¹⁹⁵ выявили, что баллы по спектру интернализации прогнозировали такие исходы, как возникновение в будущем симптомов интернализации (например, подавленного настроения, беспокойства, суицидальных попыток, ишемических болей в грудной клетке и язвенной болезни). Более того, прогноз этих результатов, базирующийся на дименсиональном спектре интернализации, был гораздо более точным, чем при использовании концептуализации различных расстройств интернализации на основе DSM (например, БДР, генерализованное тревожное расстройство), тем самым демонстрируя доказательства пользы этого подхода в определении тяжести в связи с такими важными исходами, как риск суицида и проблемы с физическим здоровьем¹⁹⁵.

В отношении других форм психопатологии Krueger и соавт.¹⁹⁶ представили доказательства того, что симптомы и поведение, определяющие личность, и расстройства употребления психоактивных веществ могут быть охвачены дименсией экстернализации. В других исследованиях также поддерживается идея наличия скрытой дименсии экстернализации, которая объясняет, почему асоциальное поведение (различные противоправные модели поведения), черты характера (импульсивность, черствость) и проблемы, связанные с употреблением психоактивных веществ, часто коморбидны^{191,197}. Carragher и соавт.¹⁹⁷ представили результаты, свидетельствующие о том, что соответствие критериям некоторых расстройств (например, кокаиновой зависимости) определяет более высокое положение и тяжесть на дименсии экстернализации, чем соответствие критериям других «менее тяжелых» расстройств (например, никотиновой и алкогольной зависимости). Вероятно, схожим образом перекрытие таких расстройств, как шизофрения и шизотипическое расстройство личности, отражается спектром расстройств мышления^{191,198}. Положение в этом спектре было связано с нарушениями функционирования и течением болезни¹⁹⁸.

В будущих исследованиях важно определить, как степень тяжести (т. е. насколько вероятно человек соответствует критериям различных расстройств и критериям «тяжелых» расстройств, таких как кокаиновая зависимость в случае спектра экстернализации), определенная посредством дименсий интернализации, экстернализации и расстройств мышления, прогнозирует течение расстройств и другие важные исходы, связанные с заболеваемостью и смертностью. Эти дименсии учитывают коморбидность диагнозов среди различных расстройств и, как было показано,

прогнозируют различные исходы более точно, чем диагнозы различных расстройств по DSM, предполагая преимущество данного подхода^{191,195}. В связи с этим появилась Иерархическая таксономия психопатологии (Hierarchical Taxonomy of Psychopathology – HiTOP) в качестве основной на дименсиях альтернативы DSM-5^{191,199}. Таким образом, важно определить, в какой степени эта концепция адекватно отражает психопатологическую «тяжесть», как бы она ни была определена, и полезна для исследователей и практиков.

ЗАКЛЮЧЕНИЕ

Вопрос тяжести расстройств имеет большое клиническое значение. Степень тяжести влияет на решения об объеме необходимой медицинской помощи и обращении за государственной поддержкой в связи с инвалидностью по психическим расстройствам. В амбулаторных условиях значение тяжести расстройств отражается в спорах об эффективности антидепрессантов во всем спектре тяжести депрессии, а также о том, следует ли пациентам с тяжелой депрессией предпочтительно назначать фармакотерапию, а не психотерапию.

Мы начали эту статью с серии вопросов о том, как следует концептуализировать тяжесть психопатологии. Некоторые авторы предположили, что основным показателем тяжести психического заболевания является инвалидность в связи с нарушением функционирования²⁰⁰. В DSM-5 тяжесть различных расстройств определяется по-разному. Наш обзор литературы по депрессии и РЛ показал, что исследователи используют множество способов определения тяжести. Тяжесть депрессии определялась преимущественно баллами по шкалам оценки симптомов. Однако существуют некоторые отличия в том, как оцениваются элементы (т. е. выраженность симптомов, частота симптомов, персистенция симптомов), а также некоторая вариативность в диапазоне оцениваемых симптомов при использовании различных инструментов для оценки депрессии. Независимо от конкретного способа определения тяжести симптомов, большая часть литературы по тяжести депрессии основана на свойствах симптомов. Напротив, суть патологии личности связана с ее влиянием на функционирование. Отличать крайне выраженные варианты личностных черт от нарушений функционирования непросто, поэтому нарушения функционирования легли в основу концептуализации тяжести РЛ.

Из-за того что влияние симптоматических расстройств (вроде БДР) на функционирование зачастую зависит от факторов, не связанных собственно с расстройством, таких как самоэффективность, жизнестойкость, копинг-способность, социальная поддержка, культурные и социальные ожидания, а также обязанностями, связанными с жизненной ролью и готовностью других принять эти обязанности на себя, мы утверждаем, что тяжесть подобных расстройств должна определяться независимо от нарушений функционирования. Для несогласных рассмотрим следующий сценарий: два человека имеют заболевания верхних дыхательных путей. У них одинаковое повышение температуры тела, чиханье и кашель с одинаковой частотой, одинаковый уровень выработки слизи и выделений из носовых ходов, одинаковая вирусная нагрузка. Симптомы продолжатся одинаковое количество дней. Вкратце, они имеют одинаковую выраженность, частоту и персистенцию симптомов. Тем не менее один человек отсутствует на работе в течение недели, а другой продолжает работать. Значит ли это, что человек, отсутствующий на работе, имеет более тяжелую инфекцию верхних дыхательных путей?

Можно разграничить определение тяжести на уровне расстройства и общую глобальную тяжесть заболевания. Как уже говорилось, на уровне расстройства тяжесть

должна определяться факторами, присущими расстройству. Таким образом, тяжесть депрессии должна определяться выраженностью, частотой и/или персистенцией симптомов депрессии. То же самое относится и к другим расстройствам, таким как генерализованное тревожное расстройство, посттравматическое стрессовое расстройство, мания/гипомания и тики. Тяжесть панического расстройства должна основываться на выраженности и частоте панических атак, тяжесть преждевременной эякуляции – на времени до эякуляции, тяжесть гиполипидемии – исходя из выраженности (или отсутствия) желаний, определение психогенного переживания – по частоте и выраженности эпизодов переживания и т. д. Эпизодический характер некоторых психических расстройств и симптомов создает определенные трудности для оценки. Может присутствовать вариабельность выраженности симптомов изо дня в день, так же как и персистенция симптомов может изменяться в течение дня. Частота симптомов зависит от расстройства. Слишком малое число исследований посвящено изучению валидности, оценки выраженности, частоты и персистенции симптомов.

Тяжесть, однако, можно рассматривать и с другой точки зрения – на уровне общей тяжести заболевания. Пациент с депрессией, пограничным РЛ, некоторыми тревожными расстройствами, расстройством употребления психоактивных веществ и расстройством пищевого поведения имеет тяжелое заболевание. Вероятно, будет трудно сопоставить уровни нарушений функционирования среди отдельных расстройств. Тяжесть симптомов депрессии может быть невысокой, но пациент, тем не менее, может быть тяжело болен. Неясно, как учитывать сопутствующую патологию при определении тяжести отдельных расстройств. Глобальная оценка общей тяжести заболевания была включена в DSM-III – DSM-IV, но исключена из DSM-5. Глобальная оценка тяжести заболевания может считаться похожей на комплексную оценку тяжести соматического заболевания, описанную во введении, которая используется для прогнозирования смертности в отделениях неотложной помощи и у стационарных больных. Проблема со шкалой GAF заключалась в том, что это был единый рейтинг, который требовал анализа нескольких конструкций, включая частоту симптомов, тип симптома, уровень нарушений, риск суицида, способность заботиться о себе и психоз. Из-за сложности анализа возникли проблемы с надежностью результатов²⁰¹. Возможно, способы оценки психопатологии на основе дименсий, сформулированные в HiTOP, приведут к появлению клинически содержательных и полезных подходов к характеристике общей тяжести.

В будущем в исследованиях по изучению тяжести расстройств авторы должны предполагать, с чем будет коррелировать метод оценки тяжести. Мы отметили выше, что слишком мало исследований посвящено изучению валидности оценки выраженности симптомов, их частоты и персистенции. Вопрос в том, как оценить валидность. Должна ли тяжесть быть предиктором исхода? Должна ли ее степень помогать пациентам подобрать соответствующее лечение или соответствующий объем необходимой помощи? Должна ли она прогнозировать смертность? Должна ли она отражать подлежащие патофизиологические механизмы? Должна ли она предоставлять информацию о генетическом риске? Должно ли осуществляться распределение ресурсов на уровне страховой компании или государственного финансирующего учреждения на основании тяжести?

Существует множество работ в области психиатрии, медицины и эпидемиологии, которые ссылаются на тяжесть депрессии в названии и исследуют корреляты инструментов для оценки симптомов депрессии. Но то, как лучше оценивать тяжесть, во многом не было предметом исследований. Были разработаны многочисленные шкалы,

предназначенные для измерения тяжести депрессии. Когда авторы этих шкал обсуждают причину разработки своего способа оценки, объяснение обычно нацелено на содержание и редко указывает на причину выбора конкретного подхода для оценки. Возможно, не имеет значения, как элементы масштабируются на шкале. Возможно, точное содержание шкалы также не имеет значения. Возможно, простота и клиническая практичность должны превзойти любую незначительную добавочную валидность, которой одна оценка обладает по отношению к другой.

Однако некоторые исследования указывают на обратное. Способность выявлять разницу между лекарством и плацебо может быть связана с содержанием используемого инструмента²⁰². Шкалы различаются по классификациям тяжести^{111,112,114}, и рекомендации по лечению предлагают обосновывать выбор между альтернативными вариантами лечения степенью тяжести^{1,2}. Таким образом, тяжесть расстройств имеет значение как для научных исследований, так и для клинических решений. Мы надеемся, что эта статья будет способствовать более глубокому рассмотрению и изучению вопроса о том, как наилучшим образом концептуализировать и измерять тяжесть психических расстройств.

Библиография

1. American Psychiatric Association. Practice guideline for the treatment of patients with major depressive disorder, 3rd ed. Washington: American Psychiatric Association, 2010.
2. National Collaborating Centre for Mental Health. Depression: the treatment and management of depression in adults. London: National Institute for Health and Clinical Excellence, 2009.
3. Tyrer P, Crawford M, Mulder R et al. Reclassifying personality disorders. *Lancet* 2011;377:1814-5.
4. Tyrer P, Crawford M, Mulder R et al. The rationale for the reclassification of personality disorder in the 11th revision of the International Classification of Diseases (ICD-11). *Personal Ment Health* 2011;5:246-59.
5. Minichiello E, Semerano L, Boissier MC. Time trends in the incidence, prevalence, and severity of rheumatoid arthritis: a systematic literature review. *Joint Bone Spine* 2016;83:625-30.
6. Hirai FE, Tielsch JM, Klein BE et al. Relationship between retinopathy severity, visual impairment and depression in persons with long-term type 1 diabetes. *Ophthalmic Epidemiol* 2012;19:196-203.
7. Greco A, Steca P, Pozzi R et al. Predicting depression from illness severity in cardiovascular disease patients: self-efficacy beliefs, illness perception, and perceived social support as mediators. *Int J Behav Med* 2014;21:221-9.
8. Steca P, Greco A, Monzani D et al. How does illness severity influence depression, health satisfaction and life satisfaction in patients with cardiovascular disease? The mediating role of illness perception and self-efficacy beliefs. *Psychol Health* 2013;28:765-83.
9. Pelletier R, Lavoie KL, Bacon SL et al. Depression and disease severity in patients with premature acute coronary syndrome. *Am J Med* 2014;127:87-93.
10. Carels RA. The association between disease severity, functional status, depression and daily quality of life in congestive heart failure patients. *Qual Life Res* 2004;13:63-72.
11. Snell C, Fernandes S, Bujoreanu IS et al. Depression, illness severity, and healthcare utilization in cystic fibrosis. *Pediatr Pulmonol* 2014;49:1177-81.
12. Deterding K, Gruner N, Buggisch P et al. Symptoms of anxiety and depression are frequent in patients with acute hepatitis C and are not associated with disease severity. *Eur J Gastroenterol Hepatol* 2016;28:187-92.
13. Euesden J, Matcham F, Hotopf M et al. The relationship between mental health, disease severity, and genetic risk for depression in early rheumatoid arthritis. *Psychosom Med* 2017;79:638-45.
14. Reynolds JC, Rittenberger JC, Toma C et al. Risk-adjusted outcome prediction with initial post-cardiac arrest illness severity: implications for cardiac arrest survivors being considered for early invasive strategy. *Resuscitation* 2014;85:1232-9.
15. Coppler PJ, Elmer J, Calderon L et al. Validation of the Pittsburgh Cardiac Arrest Category illness severity score. *Resuscitation* 2015;89:86-92.

16. Schaeffer JJ, Gil KM, Burchinal M et al. Depression, disease severity, and sickle cell disease. *J Behav Med* 1999;22:115-26.
17. Kim KU, Park HK, Jung HY et al. Association of depression with disease severity in patients with chronic obstructive pulmonary disease. *Lung* 2014;192:243-9.
18. van Dijk JP, Havlikova E, Rosenberger J et al. Influence of disease severity on fatigue in patients with Parkinson's disease is mainly mediated by symptoms of depression. *Eur Neurol* 2013;70:201-9.
19. Goetz CG, Tilley BC, Shaftman SR et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS UPDRS): scale presentation and clinimetric testing results. *Mov Disord* 2008;23:2129-70.
20. Bennett JA, Riegel B, Bittner V et al. Validity and reliability of the NYHA classes for measuring research outcomes in patients with cardiac disease. *Heart Lung* 2002;31:262-70.
21. Knaus WA, Zimmerman JE, Wagner DP et al. APACHE-acute physiology and chronic health evaluation: a physiologically based classification system. *Crit Care Med* 1981;9:591-7.
22. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993;270:2957-63.
23. de Groot B, de Deckere ER, Flameling R et al. Performance of illness severity scores to guide disposition of emergency department patients with severe sepsis or septic shock. *Eur J Emerg Med* 2012;19:316-22.
24. Schneider AG, Lipcsey M, Bailey M et al. Simple translational equations to compare illness severity scores in intensive care trials. *J Crit Care* 2013;28:e881-8.
25. Barry MJ, Fowler FJ Jr, O'Leary MP et al. The American Urological Association symptom index for benign prostatic hyperplasia. *J Urol* 2017;197:S189-97.
26. Folmer RL, Shi YB. SSRI use by tinnitus patients: interactions between depression and tinnitus severity. *Ear Nose Throat J* 2004;83:107-8.
27. Gros DF, Antony MM, McCabe RE et al. Frequency and severity of the symptoms of irritable bowel syndrome across the anxiety disorders and depression. *J Anxiety Disord* 2009;23:290-6.
28. Stewart WF, Lipton RB, Simon D et al. Validity of an illness severity measure for headache in a population sample of migraine sufferers. *Pain* 1999;79:291-301.
29. Baker GA, Smith DF, Jacoby A et al. Liverpool Seizure Severity Scale revisited. *Seizure* 1998;7:201-5.
30. Clark JA, Spiro A, Miller DR et al. Patient-based measures of illness severity in the Veterans Health Study. *J Ambul Care Manage* 2005;28:274-85.
31. Slade M, Powell R, Strathdee G. Current approaches to identifying the severely mentally ill. *Soc Psychiatry Psychiatr Epidemiol* 1997;32:177-84.
32. Ruggeri M, Leese M, Thornicroft G et al. Definition and prevalence of severe and persistent mental illness. *Br J Psychiatry* 2000;177:149-55.
33. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 5th ed. Arlington: American Psychiatric Association, 2013.
34. Fattori A, Neri L, Bellomo A et al. Depression severity and concentration difficulties are independently associated with HRQOL in patients with unipolar depressive disorders. *Qual Life Res* 2017;26:2459-69.
35. Goethe JW, Fischer EH, Wright JS. Severity as a key construct in depression. *J Nerv Ment Dis* 1993;181:718-24.
36. Luty SE, Joyce PR, Mulder RT et al. Social adjustment in depression: the impact of depression severity, personality, and clinic versus community sampling. *J Affect Disord* 2002;70:143-54.
37. Bradvik L, Mattisson C, Bogren M et al. Long-term suicide risk of depression in the Lundby cohort 1947-1997 – severity and gender. *Acta Psychiatr Scand* 2008;117:185-91.
38. Kessing LV. Severity of depressive episodes according to ICD-10: prediction of risk of relapse and suicide. *Br J Psychiatry* 2004;184:153-6.
39. Wang YY, Jiang NZ, Cheung EF et al. Role of depression severity and impulsivity in the relationship between hopelessness and suicidal ideation in patients with major depressive disorder. *J Affect Disord* 2015;183:83-9.
40. Katon W, Unutzer J, Russo J. Major depression: the importance of clinical characteristics and treatment response to prognosis. *Depress Anxiety* 2010;27:19-26.
41. Keller MB, Lavori PW, Mueller TI et al. Time to recovery, chronicity, and levels of psychopathology in major depression. A 5-year prospective follow-up of 431 subjects. *Arch Gen Psychiatry* 1992;49:809-16.
42. Meyers BS, Sirey JA, Bruce M et al. Predictors of early recovery from major depression among persons admitted to community-based clinics: an observational study. *Arch Gen Psychiatry* 2002;59:729-35.
43. Melartin T, Rytala H, Leskela U et al. Severity and comorbidity predict episode duration and recurrence of DSM-IV major depressive disorder. *J Clin Psychiatry* 2004;65:810-9.
44. Berent D, Zboralski K, Orzechowska A et al. Thyroid hormones association with depression severity and clinical outcome in patients with major depressive disorder. *Mol Biol Rep* 2014;41:2419-25.
45. de Diego-Adelino J, Pires P, Gomez-Anson B et al. Microstructural whitematter abnormalities associated with treatment resistance, severity and duration of illness in major depression. *Psychol Med* 2014;44:1171-82.
46. Zimmerman M, Coryell W, Pfohl B. The validity of the dexamethasone suppression test as a marker for endogenous depression. *Arch Gen Psychiatry* 1986;43:347-55.
47. Lux V, Aggen SH, Kendler KS. The DSM-IV definition of severity of major depression: inter-relationship and validity. *Psychol Med* 2010;40:1691-701.
48. Faravelli C, Servi P, Arends J et al. Number of symptoms, quantification, and qualification of depression. *Compr Psychiatry* 1996;37:307-15.
49. Guy W. ECDEU assessment manual for psychopharmacology. Rockville: National Institute of Mental Health, 1976.
50. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed. Washington: American Psychiatric Association, 1994.
51. Kitamura T, Nakagawa Y, Machizawa S. Grading depression severity by symptom scores: is it a valid method for subclassifying depressive disorders? *Compr Psychiatry* 1993;34:280-3.
52. Kessler RC, Zhao S, Blazer DG et al. Prevalence, correlates, and course of minor depression and major depression in the National Comorbidity Survey. *J Affect Disord* 1997;45:19-30.
53. Wakefield JC, Schmitz MF. Severity of complicated versus uncomplicated subthreshold depression: new evidence on the "monotonicity thesis" from the National Comorbidity Survey. *J Affect Disord* 2017;212:101-9.
54. Wakefield JC, Schmitz MF. Symptom quality versus quantity in judging prognosis: using NESARC predictive validators to locate uncomplicated major depression on the number-of-symptoms severity continuum. *J Affect Disord* 2017;208:325-9.
55. World Health Organization. International classification of diseases and related health problems, 10th revision. Geneva: World Health Organization, 2016.
56. Hiller W, Dichtl G, Hecht H et al. Evaluating the new ICD-10 categories of depressive episode and recurrent depressive disorder. *J Affect Disord* 1994;31:49-60.
57. Montgomery S. Are the ICD-10 or DSM-5 diagnostic systems able to define those who will benefit from treatment for depression? *CNS Spectr* 2016;21:283-8.
58. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56-62.
59. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:382-9.
60. Nemeroff CB. The burden of severe depression: a review of diagnostic challenges and treatment alternatives. *J Psychiatr Res* 2007;41:189-206.
61. Sonawalla SB, Fava M. Severe depression: is there a best approach? *CNS Drugs* 2001;15:765-76.
62. Williams JB, Kobak KA, Bech P et al. The GRID-HAMD: standardization of the Hamilton Depression Rating Scale. *Int Clin Psychopharmacol* 2008;23:120-9.
63. Parker G, Hadzi-Pavlovic D, Sengoz A et al. A brief self-report depression measure assessing mood state and social impairment. *J Affect Disord* 1994;30:133-42.
64. Pilkonis PA, Choi SW, Reise SP et al. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS): depression, anxiety, and anger. *Assessment* 2011;18:263-83.
65. Balsamo M, Giampaglia G, Saggino A. Building a new Rasch-based self-report inventory of depression. *Neuropsychiatr Dis Treat* 2014;10:153-65.
66. Vaccarino AL, Evans KR, Kalali AH et al. The Depression Inventory Development Workgroup: a collaborative, empirically driven initiative

- to develop a new assessment tool for major depressive disorder. *Innov Clin Neurosci* 2016;13:20-31.
67. Zimmerman M, Posternak M, Friedman M et al. Which factors influence psychiatrists' selection of an antidepressant? *Am J Psychiatry* 2004;161:1285-9.
 68. Cheung HN, Power MJ. The development of a new multidimensional depression assessment scale: preliminary results. *Clin Psychol Psychother* 2012;19:170-8.
 69. Licht RW, Qvitzau S, Allerup P et al. Validation of the Bech-Rafaelsen Melancholia Scale and the Hamilton Depression Scale in patients with major depression; is the total score a valid measure of illness severity? *Acta Psychiatr Scand* 2005;111:144-9.
 70. Santor D, Coyne J. Examining symptom expression as a function of symptom severity: item performance on the Hamilton Rating Scale for Depression. *Psychol Assess* 2001;13:127-39.
 71. Kroenke K, Spitzer R, Williams J. The PHQ-9. Validity of a brief depression severity measure. *J Gen Int Med* 2001;16:606-13.
 72. Downey L, Hayduk LA, Curtis JR et al. Measuring depression-severity in critically ill patients' families with the Patient Health Questionnaire (PHQ): tests for unidimensionality and longitudinal measurement invariance, with implications for CONSORT. *J Pain Sympt Manage* 2016;51:938-46.
 73. Fischer HF, Tritt K, Klapp BF et al. How to compare scores from different depression scales: equating the Patient Health Questionnaire (PHQ) and the ICD-10-Symptom Rating (ISR) using item response theory. *Int J Methods Psychiatr Res* 2011;20:203-14.
 74. Adler M, Hetta J, Isacson G et al. An item response theory evaluation of three depression assessment instruments in a clinical sample. *BMC Med Res Methodol* 2012;12:84.
 75. Barthel D, Barkmann C, Ehrhardt S et al. Screening for depression in pregnant women from Cote d'Ivoire and Ghana: psychometric properties of the Patient Health Questionnaire-9. *J Affect Disord* 2015;187:232-40.
 76. Pedersen SS, Mathiasen K, Christensen KB et al. Psychometric analysis of the Patient Health Questionnaire in Danish patients with an implantable cardioverter defibrillator (The DEFIB-WOMEN study). *J Psychosom Res* 2016;90:105-12.
 77. Umegaki Y, Todo N. Psychometric properties of the Japanese CES-D, SDS, and PHQ-9 depression scales in university students. *Psychol Assess* 2017;29:354-9.
 78. Zhong Q, Gelaye B, Fann JR et al. Cross-cultural validity of the Spanish version of PHQ-9 among pregnant Peruvian women: a Rasch item response theory analysis. *J Affect Disord* 2014;158:148-53.
 79. Zimmerman M, Chelminski I, McGlinchey JB et al. A clinically useful depression outcome scale. *Compr Psychiatry* 2008;49:131-40.
 80. Zimmerman M, Galione J, Attiullah N et al. Depressed patients perspectives of two measures of outcome: the Quick Inventory of Depressive Symptomatology (QIDS) and the Remission from Depression Questionnaire (RDQ). *Ann Clin Psychiatry* 2011;23:208-12.
 81. Bentley KH, Gallagher MW, Carl JR et al. Development and validation of the Overall Depression Severity and Impairment Scale. *Psychol Assess* 2014;26:815-30.
 82. Bagby RM, Ryder AG, Schuller DR et al. The Hamilton Depression Rating Scale: has the gold standard become a lead weight? *Am J Psychiatry* 2004;161:2163-77.
 83. Zimmerman M, Posternak M, Chelminski I. Is it time to replace the Hamilton Depression Rating Scale as the primary outcome measure in treatment studies of depression? *J Clin Psychopharmacol* 2005;25:105-10.
 84. Tiplady B. A self-rating scale for depression designed to be sensitive to change. *Neuropharmacology* 1980;19:1211-2.
 85. Rush A, Trivedi M, Ibrahim H et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry* 2003;54:573-83.
 86. Olsen LR, Jensen DV, Noerholm V et al. The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. *Psychol Med* 2003;33:351-6.
 87. Rush AJ, Gullion CM, Basco MR et al. The Inventory of Depressive Symptomatology (IDS). *Psychol Med* 1996;26:477-86.
 88. Moller HJ. Methodological aspects in the assessment of severity of depression by the Hamilton Depression Scale. *Eur Arch Psychiatry Clin Neurosci* 2001;251(Suppl. 2):13-20.
 89. DeRubeis RJ, Gelfand LA, Tang TZ et al. Medications versus cognitive behavior therapy for severely depressed outpatients: mega-analysis of four randomized comparisons. *Am J Psychiatry* 1999;156:1007-13.
 90. Gibbons R, Hur K, Brown C et al. Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. *Arch Gen Psychiatry* 2012;69:572-9.
 91. Elkin I, Gibbons R, Shea M et al. Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *J Consult Clin Psychol* 1995;63:841-7.
 92. Kirsch I, Deacon BJ, Huedo-Medina TB et al. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* 2008;5:260-8.
 93. Rush AJ, First MB, Blacker D. *Handbook of psychiatric measures*, 2nd ed. Washington: American Psychiatric Publishing, 2008.
 94. Khan A, Leventhal RM, Khan SR et al. Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. *J Clin Psychopharmacol* 2002;22:40-5.
 95. Fournier JC, DeRubeis RJ, Hollon SD et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010;303:47-53.
 96. Dunner D, Lipschitz A, Pitts C et al. Efficacy and tolerability of controlled-release paroxetine in the treatment of severe depression: post hoc analysis of pooled data from a subset of subjects in four double-blind clinical trials. *Clin Ther* 2005;27:1901-11.
 97. Kasper S. Efficacy of antidepressants in the treatment of severe depression: the place of mirtazapine. *J Clin Psychopharmacol* 1997;17(Suppl. 1):19S-28S.
 98. Montgomery S, Ferguson J, Schwartz G. The antidepressant efficacy of reboxetine in patients with severe depression. *J Clin Psychopharmacol* 2003;23:45-50.
 99. Schmitt AB, Bauer M, Volz HP et al. Differential effects of venlafaxine in the treatment of major depressive disorder according to baseline severity. *Eur Arch Psychiatry Clin Neurosci* 2009;259:329-39.
 100. Shelton RC, Prakash A, Mallinckrodt CH et al. Patterns of depressive symptom response in duloxetine-treated outpatients with mild, moderate or more severe depression. *Int J Clin Pract* 2007;61:1337-48.
 101. Versiani M, Moreno R, Ramakers-van Moorsel C et al. Comparison of the effects of mirtazapine and fluoxetine in severely depressed patients. *CNS Drugs* 2005;19:137-46.
 102. Hirschfeld R. Efficacy of SSRIs and newer antidepressants in severe depression: comparison with TCAs. *J Clin Psychiatry* 1999;60:326-35.
 103. Montgomery S, Lecrubier Y. Is severe depression a separate indication? *Eur Neuropsychopharmacol* 1999;9:259-64.
 104. Schatzberg AF. Antidepressant effectiveness in severe depression and melancholia. *J Clin Psychiatry* 1999;60(Suppl. 4):14-21.
 105. Endicott J, Cohen J, Nee J et al. Hamilton depression rating scale. *Arch Gen Psychiatry* 1981;38:98-103.
 106. Kearns NP, Cruickshank CA, McGuigan KJ et al. A comparison of depression rating scales. *Br J Psychiatry* 1982;141:45-9.
 107. Cameron IM, Cardy A, Crawford JR et al. Measuring depression severity in general practice: discriminatory performance of the PHQ-9, HADS-D, and BDI-II. *Br J Gen Pract* 2011;61:e419-26.
 108. Zimmerman M, Martinez JH, Young D et al. Severity classification on the Hamilton Depression Rating Scale. *J Affect Disord* 2013;150:384-8.
 109. Zigmund AS, Snaith RP. *The Hospital Anxiety and Depression Scale*. *Acta Psychiatr Scand* 1983;67:361-70.
 110. Cameron IM, Crawford JR, Lawton K et al. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *Br J Gen Pract* 2008;58:32-6.
 111. Hansson M, Chotai J, Nordstrom A et al. Comparison of two self-rating scales to detect depression: HADS and PHQ-9. *Br J Gen Pract* 2009;59: e283-8.
 112. Reddy P, Philpot B, Ford D et al. Identification of depression in diabetes: the efficacy of PHQ-9 and HADS-D. *Br J Gen Pract* 2010;60:e239-45.
 113. Beck A, Steer R, Brown G. *The Beck Depression Inventory*, 2nd ed. San Antonio: The Psychological Corporation, 1996.
 114. Zimmerman M, Martinez J, Friedman M et al. How can we use depression severity to guide treatment selection when measures of depression categorize patients differently? *J Clin Psychiatry* 2012;73:1287-91.
 115. Ruscio J, Zimmerman M, McGlinchey JB et al. Diagnosing major depressive disorder XI: a taxometric investigation of the structure underlying DSM-IV symptoms. *J Nerv Ment Dis* 2007;195:10-9.

116. van der Lem R, van der Wee NJ, van Veen T et al. The generalizability of antidepressant efficacy trials to routine psychiatric out-patient practice. *Psychol Med* 2011;41:1353-63.
117. Bielski RJ, Friedel RO. Prediction of tricyclic antidepressant response: a critical review. *Arch Gen Psychiatry* 1976;33:1479-89.
118. Grammer GG, Kuhle AR, Clark CC et al. Severity of depression predicts remission rates using transcranial magnetic stimulation. *Front Psychiatry* 2015;6:114.
119. Jones NP, Siegle GJ, Thase ME. Effects of rumination and initial severity on remission to cognitive therapy for depression. *Cogn Ther Res* 2008;32:591-604.
120. Lisanby SH, Husain MM, Rosenquist PB et al. Daily left prefrontal repetitive transcranial magnetic stimulation in the acute treatment of major depression: clinical predictors of outcome in a multisite, randomized controlled clinical trial. *Neuropsychopharmacology* 2009;34:522-34.
121. Bower P, Kontopantelis E, Sutton A et al. Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. *BMJ* 2013;346:f540.
122. Sugawara Y, Higuchi H, Yoshida K et al. Response rate obtained using milnacipran depending on the severity of depression in the treatment of major depressive patients. *Clin Neuropharmacol* 2006;29:6-9.
123. Friedman ES, Davis LL, Zisook S et al. Baseline depression severity as a predictor of single and combination antidepressant treatment outcome: results from the CO-MED trial. *Eur Neuropsychopharmacol* 2012;22:183-99.
124. Kennedy S, Andersen H, Lam R. Efficacy of escitalopram in the treatment of major depressive disorder compared with conventional selective serotonin reuptake inhibitors and venlafaxine SR: a meta-analysis. *J Psychiatry Neurosci* 2006;31:122-31.
125. Angst J, Amrein R, Stahl M. Moclobemide and tricyclic antidepressants in severe depression: meta-analysis and prospective studies. *J Clin Psychopharmacol* 1995;15:16S-23S.
126. Wiles NJ, Mulligan J, Peters TJ et al. Severity of depression and response to antidepressants: GENPOD randomised controlled trial. *Br J Psychiatry* 2011;200:130-6.
127. Kilts CD, Wade AG, Andersen HF et al. Baseline severity of depression predicts antidepressant drug response relative to escitalopram. *Exp Opin Pharmacother* 2009;10:927-36.
128. Khan A, Sambunaris A, Edwards J et al. Vilazodone in the treatment of major depressive disorder: efficacy across symptoms and severity of depression. *Int Clin Psychopharmacol* 2014;29:86-92.
129. Mosca D, Zhang M, Prieto R et al. Efficacy of desvenlafaxine compared with placebo in major depressive disorder patients by age group and severity of depression at baseline. *J Clin Psychopharmacol* 2017;37:182-92.
130. Driessen E, Cuijpers P, Hollon SD et al. Does pretreatment severity moderate the efficacy of psychological treatment of adult outpatient depression? A meta-analysis. *J Consult Clin Psychol* 2010;78:668-80.
131. Furukawa TA, Weitz ES, Tanaka S et al. Initial severity of depression and efficacy of cognitive-behavioural therapy: individual-participant data metaanalysis of pill-placebo-controlled trials. *Br J Psychiatry* 2017;210:190-6.
132. Guelfi JD, Ansseau M, Timmerman L et al. Mirtazapine versus venlafaxine in hospitalized severely depressed patients with melancholic features. *J Clin Psychopharmacol* 2001;21:425-31.
133. Henkel V, Seemuller F, Obermeier M et al. Relationship between baseline severity of depression and antidepressant treatment outcome. *Pharmacopsychiatry* 2011;44:27-32.
134. Hirschfeld RM, Russell JM, Delgado PL et al. Predictors of response to acute treatment of chronic and double depression with sertraline or imipramine. *J Clin Psychiatry* 1998;59:669-75.
135. Madhoo M, Levine SZ. Initial severity effects on residual symptoms in response and remission: a STAR*D study during and after failed citalopram treatment. *J Clin Psychopharmacol* 2015;35:450-3.
136. Souery D, Oswald P, Massat I et al. Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. *J Clin Psychiatry* 2007;68:1062-70.
137. Brown C, Schulberg HC, Prigerson HG. Factors associated with symptomatic improvement and recovery from major depression in primary care patients. *Gen Hosp Psychiatry* 2000;22:242-50.
138. Enns MW, Cox BJ. Psychosocial and clinical predictors of symptom persistence vs. remission in major depressive disorder. *Can J Psychiatry* 2005;50:769-77.
139. Sargeant JK, Bruce ML, Florio LP et al. Factors associated with 1-year outcome of major depression in the community. *Arch Gen Psychiatry* 1990;47:519-26.
140. Turner EH, Matthews AM, Linardatos E et al. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;358:252-60.
141. Rabinowitz J, Werbeloff N, Mandel FS et al. Initial depression severity and response to antidepressants vs. placebo: patient-level data analysis from 34 randomised controlled trials. *Br J Psychiatry* 2016;209:427-8.
142. Elkin I, Shea M, Watkins J et al. NIMH treatment of depression collaborative research program: general effectiveness of treatments. *Arch Gen Psychiatry* 1989;46:971-82.
143. Cuijpers P, van Straten A, van Oppen P et al. Are psychological and pharmacologic interventions equally effective in the treatment of adult depressive disorders? A meta-analysis of comparative studies. *J Clin Psychiatry* 2008;69:1675-85.
144. Weitz ES, Hollon SD, Twisk J et al. Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs. pharmacotherapy: an individual patient data meta-analysis. *JAMA Psychiatry* 2015;72:1102-9.
145. Links PS, Eynan R. The relationship between personality disorders and Axis I psychopathology: deconstructing comorbidity. *Annu Rev Clin Psychol* 2013;9:529-54.
146. Lenzenweger MF, Lane MC, Loranger AW et al. DSM-IV personality disorders in the National Comorbidity Survey Replication. *Biol Psychiatry* 2007;62:553-64.
147. Bateman AW, Fonagy P. The effectiveness of partial hospitalization in the treatment of borderline personality disorder – a randomised controlled trial. *Am J Psychiatry* 1999;156:1563-9.
148. National Collaborating Centre for Mental Health. *Borderline personality disorder: recognition and management*. London: Department of Health, 2009.
149. Blum N, St John D, Pfohl B et al. Systems Training for Emotional Predictability and Problem Solving (STEPPS) for outpatients with borderline personality disorder: a randomized controlled trial and 1-year follow-up. *Am J Psychiatry* 2008;165:468-78.
150. Zanarini MC, Yong L, Frankenburg FR et al. Severity of childhood sexual abuse and its relationship to severity of borderline psychopathology and psychosocial impairment among borderline inpatients. *J Nerv Ment Dis* 2002;190:381-7.
151. Giesen-Bloo JH, Wouters LM, Schouten E et al. The borderline personality disorder severity index-IV: psychometric evaluation and dimensional structure. *Pers Ind Diff* 2010;49:136-41.
152. Yang M, Coid J, Tyrer P. Personality pathology recorded by severity: national survey. *Br J Psychiatry* 2010;197:193-9.
153. Tyrer P, Alexander MS, Cicchetti D et al. Reliability of a schedule for rating personality disorders. *Br J Psychiatry* 1979;135:168-74.
154. Dolan B, Evans C, Norton K. Multiple axis-II diagnoses of personality disorder. *Br J Psychiatry* 1995;166:107-12.
155. Oldham JM, Skodol AE, Kellman HD et al. Diagnosis of DSM-III-R personality disorders by two structured interviews: patterns of comorbidity. *Am J Psychiatry* 1992;149:213-20.
156. Zimmerman M, Galione JN, Chelminski I et al. Does the diagnosis of multiple Axis II disorders have clinical significance? *Ann Clin Psychiatry* 2012;24:195-201.
157. Tyrer P, Johnson T. Establishing the severity of personality disorder. *Am J Psychiatry* 1996;153:1593-7.
158. Asnaani A, Chelminski I, Young D et al. Heterogeneity of borderline personality disorder: do the number of criteria met make a difference? *J Pers Disord* 2007;21:615-25.
159. Kernberg OF, Caligor E. A psychoanalytic theory of personality disorders. In: Clarkin JF, Lenzenweger MF (eds). *Major theories of personality disorder*. New York: Guilford, 2005:114-56.
160. Paris J. Dimensional diagnosis and the DSM-5. *J Clin Psychiatry* 2005;72:1340.
161. Moran P. Dangerous severe personality disorder – bad tidings from the UK. *Int J Soc Psychiatry* 2001;48:6-10.
162. Maden A, Tyrer P. Dangerous and severe personality disorders: a new personality concept from the United Kingdom. *J Pers Disord* 2003;17:489-96.
163. Tyrer P, Cooper S, Rutter D et al. The assessment of dangerous and severe personality disorder: lessons from a randomised controlled trial linked to qualitative analysis. *Forensic Psychol Psychiatry* 2009;20:132-46.

164. Berghuis H, Kamphuis JH, Verheul R. Specific personality traits and general personality dysfunction as predictors of the presence and severity of personality disorders in a clinical sample. *J Pers Assess* 2014;96:410-6.
165. Crawford MJ, Koldobsky N, Mulder R et al. Classifying personality disorder according to severity. *J Pers Disord* 2011;25:321-30.
166. Livesley WJ. Practical management of personality disorders. New York: Guilford, 2003.
167. Widiger TA, Costa PT, McCrae RR. A proposal for Axis II: diagnosing personality disorders using the five factor model. In: Costa PT, Widiger TA (eds). *Personality disorders and the five factor model of personality*. Washington: American Psychological Association, 2002:431-52.
168. Widiger TA, Mullins-Sweatt SN. Five-factor model of personality disorder: a proposal for DSM-V. *Annu Rev Clin Psychol* 2009;5:197-220.
169. Widiger TA, Trull TJ. Plate tectonics in the classification of personality disorder: shifting to a dimensional model. *Am Psychol* 2007;62:71-83.
170. Livesley W, Schroeder M, Jackson D et al. Categorical distinctions in the study of personality disorder: implications for classification. *J Abnorm Psychol* 1994;103:6-17.
171. Parker G, Hadzi-Pavlovic D, Both L et al. Measuring disordered personality functioning: to love and to work revisited. *Acta Psychiatr Scand* 2004;110:230-9.
172. Trull TJ. Dimensional models of personality disorder: coverage and cutoffs. *J Pers Disord* 2005;19:262-82.
173. Verheul R, Andrea H, Berghout CC et al. Severity Indices of Personality Problems (SIPP-118): development, factor structure, reliability, and validity. *Psychol Assess* 2008;20:23-34.
174. Livesley WJ. Suggestions for a framework for an empirically based classification of personality disorder. *Can J Psychiatry* 1998;43:137-47.
175. Ro E, Clark LA. Interrelations between psychosocial functioning and adaptive- and maladaptive-range personality traits. *J Abnorm Psychol* 2013;122:822-35.
176. Krueger RF, Derringer J, Markon KE et al. Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychol Med* 2012;42:1879-90.
177. Gunderson J, Zanarini MC. Commentary: deceptively simple – or radical shift? *Pers Ment Health* 2011;5:260-2.
178. Kelly BD, Nur UA, Tyrer P et al. Impact of severity of personality disorder on the outcome of depression. *Eur Psychiatry* 2009;24:322-6.
179. Livesley WJ. *General Assessment of Personality Disorder (GAPD)*. Vancouver: University of British Columbia, 2006.
180. Berghuis H, Kamphuis JH, Verheul R. Core features of personality disorder: differentiating general personality dysfunctioning from personality traits. *J Pers Disord* 2012;26:704-16.
181. Berghuis H, Kamphuis JH, Verheul R et al. The General Assessment of Personality Disorder (GAPD) as an instrument for assessing the core features of personality disorders. *Clin Psychol Psychother* 2013;20:544-57.
182. Arntz A, van den Hoorn M, Cornelis J et al. Reliability and validity of the borderline personality disorder severity index. *J Pers Disord* 2003;17:45-59.
183. Clarkin JF, Levy KN, Lenzenweger MF et al. Evaluating three treatments for borderline personality disorder: a multiwave study. *Am J Psychiatry* 2007;164:922-8.
184. Linehan MM, Comtois KA, Murray AM et al. Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs. therapy by experts for suicidal behaviors and borderline personality disorder. *Arch Gen Psychiatry* 2006;63:757-66.
185. Zanarini MC, Vujanovic AA, Parachini EA et al. Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J Pers Disord* 2003;17:233-42.
186. Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt Impulsiveness Scale. *J Clin Psychol* 1995;51:768-74.
187. Doering S, Horz S, Rentrop M et al. Transference-focused psychotherapy v. treatment by community psychotherapists for borderline personality disorder: randomised controlled trial. *Br J Psychiatry* 2010;196:389-95.
188. McMMain SF, Links PS, Gnam WH et al. A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *Am J Psychiatry* 2009;166:1365-74.
189. Diamond D, Yeomans FE, Stern B et al. Transference focused psychotherapy for patients with comorbid narcissistic and borderline personality disorder. *Psychoanal Inq* 2013;33:527-51.
190. Messina NP, Wish ED, Hoffman JA et al. Antisocial personality disorder and treatment outcomes. *Am J Drug Alcohol Abuse* 2002;28:197-212.
191. Kotov R, Krueger RF, Watson D et al. The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *J Abnorm Psychol* 2017;126:454-77.
192. Stanton K, Rozek DC, Stasik-O'Brien SM et al. A transdiagnostic approach to examining the incremental predictive power of emotion regulation and basic personality dimensions. *J Abnorm Psychol* 2016;125:960-75.
193. Zimmerman M. A review of 20 years of research on overdiagnosis and underdiagnosis in the Rhode Island Methods to Improve Diagnostic Assessment and Services (MIDAS) project. *Can J Psychiatry* 2016;61:71-9.
194. Krueger RF, Finger MS. Using item response theory to understand comorbidity among anxiety and unipolar mood disorders. *Psychol Assess* 2001;13:140-51.
195. Eaton NR, Krueger RF, Markon KE et al. The structure and predictive validity of the internalizing disorders. *J Abnorm Psychol* 2013;122:86-92.
196. Krueger RF, Markon KE, Patrick CJ et al. Externalizing psychopathology in adulthood: a dimensional-spectrum conceptualization and its implications for DSM-V. *J Abnorm Psychol* 2005;114:537-50.
197. Carragher N, Krueger RF, Eaton NR et al. ADHD and the externalizing spectrum: direct comparison of categorical, continuous, and hybrid models of liability in a nationally representative sample. *Soc Psychiatry Psychiatr Epidemiol* 2014;49:1307-17.
198. Kotov R, Chang SW, Fochtmann LJ et al. Schizophrenia in the internalizing-externalizing framework: a third dimension? *Schizophr Bull* 2011;37:1168-78.
199. Kotov R, Krueger RF, Watson D. A paradigm shift in psychiatric classification: the Hierarchical Taxonomy Of Psychopathology (HiTOP). *World Psychiatry* 2018;17:24-5.
200. Gaebel W, Zasko H, Baumann AE. The relationship between mental illness severity and stigma. *Acta Psychiatr Scand* 2006;113(Suppl. 429):41-5.
201. Grootenboer EM, Giltay EJ, van der Lem R et al. Reliability and validity of the Global Assessment of Functioning Scale in clinical outpatients with depressive disorders. *J Eval Clin Pract* 2012;18:502-7.
202. Bech P, Boyer P, Germain JM et al. HAM-D17 and HAM-D6 sensitivity to change in relation to desvenlafaxine dose and baseline depression severity in major depressive disorder. *Pharmacopsychiatry* 2010;43:271-6.

DOI:10.1002/wps.20569