Современные данные об эффективности мобильных приложений для психического здоровья при симптомах депрессии и тревоги. Метаанализ 176 рандомизированных контролируемых исследований

Jake Linardon[1,2], John Torous[3], Joseph Firth[4,5], Pim Cuijpers[6,7], Mariel Messer[1], Matthew Fuller-Tyszkiewicz[1,2]

[1]School of Psychology, Deakin University, Geelong, VIC, Australia; [2]Center for Social and Early Emotional Development, Deakin University, Burwood, VIC, Australia; [3]Department of Psychiatry, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA; [4]Division of Psychology and Mental Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK; [5]Greater Manchester Mental Health NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK; [6]Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [7]International Institute for Psychotherapy, Babes-Bolyai University, Cluj-Napoca, Romania

Психиатрическая помощь, доступная при депрессии и тревоге, в последнее время претерпела крупную технологическую революцию вместе с возрастающим интересом к приложениям для смартфонов как к масштабируемому инструменту для лечения этих состояний. С тех пор как последний всесторонний метаанализ в 2019 году выявил положительное, но изменчивое влияние приложений на симптомы депрессии и тревоги, было проведено более 100 новых рандомизированных контролируемых исследований (РКИ). Мы провели обновленный метаанализ с целью предоставить более точные оценки эффектов, количественно определить обобщаемость данной информационной базы и понять, влияют ли основные характеристики приложения и исследования на величину эффекта. В исследование были включены 176 РКИ, направленные на лечение симптомов депрессии или тревоги. Приложения в целом оказали значимое, хотя и небольшое влияние на симптомы депрессии (N=33,567, g=0,28, p<0,001; число, больных, которых необходимо лечить, ЧБНЛ=11,5) и генерализованной тревожности (N=22,394, g=0,26, p<0,001, ЧБНЛ=12,4) по сравнению с контрольными группами. Эти эффекты были устойчивыми при различных наблюдениях, а также после исключения исследований с небольшими выборками и высоким риском систематической ошибки. При посттестировании в приложении наблюдалась меньшая вариабельность результатов по сравнению с контрольными условиями (отношение дисперсии, RoV = -0,14, 95% ДИ: от -0,24 до -0,05 для депрессивных симптомов; RoV = -0,21, 95% ДИ: от -0,31 до -0,12 для симптомов генерализованной тревоги). Величина эффекта для депрессии была значительно больше, когда приложения включали характеристики когнитивно-поведенческой терапии (КПТ) или технологию чат-ботов. Величина эффекта для тревоги была значительно больше, когда в исследованиях в качестве основной цели рассматривалась генерализованная тревожность и применялось приложение КПТ или приложение с функцией мониторинга настроения. Были обнаружены доказательства умеренного влияния приложений на симптомы социальной тревожности (g=0,52) и обсессивно-компульсивные симптомы (g=0,51), небольшой эффект на симптомы посттравматического стресса (g=0,12), большое влияние на симптомы акрофобии (g=0,90), а также незначительное негативное влияние на симптомы паники (g = -0,12), хотя к этим результатам следует относиться с осторожностью, поскольку большинство исследований имели высокий риск систематической ошибки и основывались на небольших размерах выборки. Мы пришли к выводу, что приложения в целом оказывают небольшое, но существенное влияние на симптомы депрессии и генерализованной тревожности и что специфические особенности приложений, такие как функции КПТ или мониторинга настроения, а также технология чат-ботов, связаны с наибольшим эффектом.

Ключевые слова: Приложения для смартфонов, депрессия, генерализованная тревожность, социальная тревожность, посттравматический стресс, паника, когнитивно-поведенческая терапия, мониторинг настроения, технология чат-ботов

Linardon J, Torous J, Firth J et al. Current evidence on the efficacy of mental health 139 smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. *World Psychiatry*. 2024;23:139–149. doi:10.1002/wps.21183

Депрессивные и тревожные расстройства являются распространенными психическими состояниями, связанными со значимым бременем болезни, серьезными экономическими издержками, преждевременной смертностью и низким качеством жизни пострадавших лиц и их родственников[1,2]. Необходимы широкодоступные методы лечения для уменьшения этих последствий. Как психотерапия, так и фармакотерапия могут эффективно лечить симптомы депрессии и тревожности[7]. Однако существует множество препятствий, которые мешают людям получить доступ к этим традиционным формам лечения, включая ограниченное количество подготовленных психиатров и психологов, высокую стоимость лечения, стигматизацию, связанную с обращением за помощью, и низкую грамотность в области психического здоровья[8,9].

За последние два десятилетия психиатрическая помощь, доступная при депрессии и тревоге, претерпела крупную технологическую революцию. Эмпирически подтвержденные компоненты психологических вмешательств, которые ранее проводились исключительно лично, теперь предоставляются с помощью недорогих, частных и масштабируемых цифровых инструментов[10,11].

Приложения для смартфонов являются одной из форм предоставления цифрового лечения, которой уделяется значительное внимание. Смартфоны находятся в числе наиболее быстро внедряемых технологических инноваций в новейшей истории. Более 6,5 миллиарда человек владеют смартфонами, которые, как правило, проверяются многократно в течение дня и всегда находятся на расстоянии вытянутой руки[12]. Таким образом, доступ к лечебному контенту, предоставляемому через приложение, возможен в любое время и в любом месте, что позволяет пользователям в моменты необходимости практиковать те важнейшие терапевтические навыки, которые требуются для предотвращения появления или обострения симптомов[8]. Цифровые системы мониторинга и сложные алгоритмы машинного обучения также позволяют регулярно обновлять леченый контент и персонализировать его в соответствии с потребностями пользователя на основе данных, собираемых пассивно (например, координаты глобальной системы позиционирования для определения социальных детерминант здоровья) и активно (например, отслеживание симптомов)[13,14].

Потенциал приложений для лечения симптомов депрессии и тревоги привлекает все больший интерес среди пациентов, клиницистов, технологических компаний и регулирующих органов здравоохранения. Однако существуют риски, связанные с приложениями для лечения депрессии

и тревожности, такие как нарушения конфиденциальности; возможный легкий доступ к неэффективному, неточному или потенциально вредному контенту[15,16]; низкий уровень вовлеченности[17,18] и исключение потенциальной терапевтической составляющей, представленной личными отношениями между врачом и пациентом. Все это, в сочетании с тем фактом, что приложения для поддержания психического здоровья при депрессии и тревоге являются одними из наиболее широко предлагаемых и загружаемых категорий медицинских приложений[19,20], порождает обязанность предоставлять общественности актуальную информацию о доказательной базе, поддерживающей их использование[21].

В метаанализе 2019 года Linardon et al.[13] обнаружили, что приложения превосходят контрольные условия в уменьшении симптомов депрессии (g=0,28) и генерализованной тревожности (g=0,30) на основе 41 и 28 исследований соответственно. Этот метаанализ также выявил ранние доказательства эффективности приложений при социальной тревожности, но не при посттравматическом стрессе или панических симптомах (≤6 исследований). Была отмечена неоднородность в эффективности, хотя было выявлено несколько надежных модификаторов эффекта, что, возможно, отражает относительно небольшое количество доступных исследований[22] и фокус на однофакторных модераторных эффектах, а не на более сложных многофакторных моделях модераторов.

Число исследований, тестирующих приложения для психического здоровья при симптомах депрессии и тревоги, растет в геометрической прогрессии, предоставляя все больше возможностей выяснить, способствовали ли недавние инновации в области цифрового здравоохранения повышению эффективности, а также изучить индивидуальное и комбинированное влияние модераторов на эффективность лечения. С 2019 года было опубликовано более 100 рандомизированных контролируемых исследований (РКИ), часть из которых включает большие размеры выборки, использует надежные условия сравнения и имеет меньший риск систематической ошибки. Каждый из этих элементов был отмечен в качестве критических характеристик, которые отсутствуют в данной области, в соответствии с выводами, сделанными авторами более ранних метаанализов[13,21,23]. Кроме того, большое количество доступных исследований теперь позволяет проводить более точный, сложный и адекватно подкрепленный анализ тех характеристик приложений и исследований, которые могут быть связаны с величиной эффекта.

В свете ограничений предыдущих обзоров и в связи с дальнейшим расширением интереса к исследованиям приложений для психического здоровья мы провели обновленный метаанализ, изучающий влияние приложений для психического здоровья на симптомы депрессии и тревоги. В дополнение к типичному объединению средних различий между вмешательством и контрольными условиями мы также исследовали групповые различия в вариабельности результатов, чтобы оценить потенциальную обобщаемость эффектов этих приложений. Как утверждалось ранее[24,25], если вмешательство оказывает изменчивый эффект на участников, подобное может наблюдаться за счет большей вариабельности пост-тестового среднего значения для группы вмешательства по сравнению с контрольной группой. Наконец, в свете очевидной неоднородности величины эффекта из предыдущих обзоров[13,21], мы попытались определить модераторы (в частности, основные характеристики приложения, исследуемую популяцию и дизайн исследования), которые могут объяснять большую или меньшую величину эффекта, чем совокупное среднее значение.

Кроме того, целью этого метаанализа было выйти за рамки изучения отдельных модераторов, а также – впервые – оценить потенциальные комбинации и взаимодействия между предварительно выбранными переменными модератора[26]. Это поможет пролить свет на контексты, в которых определенные составляющие вмешательства могут быть наиболее эффективными, а также дополнительно охарактеризовать подгруппы людей, у которых приложения для психического здоровья вызвали наилучший ответ.

МЕТОДЫ

Идентификация и отбор испытаний

Этот обзор был проведен в соответствии с руководством «Предпочтительные элементы отчетов для систематических обзоров и метаанализов» (PRISMA)[27], и придерживался предварительно зарегистрированного протокола (CRD42023437664). Мы провели поиск (последнее обновление в июне 2023 года) в базах данных Medline, PsycINFO, Web of Science и ProQuest Database for Di ssertations, объединяющих ключевые термины, относящиеся к смартфонам, РКИ и тревоге или депрессии (см. дополнительную информацию для получения полной стратегии поиска). Мы также провели ручной поиск с помощью соответствующих обзоров[13,14,21] и библиографических списков включенных исследований для выявления любых потенциально приемлемых исследований, не охваченных первичным поиском.

Мы включили РКИ, в которых исследовалось влияние автономного приложения для психического здоровья на базе смартфона в сравнении с контрольными условиями (например, список ожидания, плацебо, информационные ресурсы) или активное сравнение (например, очное лечение) для симптомов депрессии или тревоги.

Исследования, в которых проводили одну психообразовательную или информационную сессию, предшествующую установке программы с приложением, имели право на включение. Были исключены смешанные веб-программы и программы на основе приложений, а также приложения, не ориентированные – полностью или частично – на психическое здоровье (например, приложения для похудения или диеты). Были исключены вмешательства, основанные только на текстовых сообщениях. Дополнительные методы лечения также были исключены, например, когда приложения были частью более широкой программы очной психотерапии. Однако были включены пробные группы, состоящие из приложения для психического здоровья и обычного лечения, при условии, что компонент обычного лечения не состоял из структурированной программы психологического лечения (например, были разрешены исследования пациентов, продолжающих получать обычное лечение ввиду основного заболевания). Опубликованные и неопубликованные исследования имели право на включение. Если исследование не включало данные для расчета величины эффекта, с авторами связывались, и исследование исключалось, если они не предоставляли данные.

Оценка качества и извлечение данных

Мы использовали критерии из Инструментов риска систематических ошибок Кокрейновского содружества (Cochrane Collaboration Risk of Bias tool)[28] для оценки риска систематической ошибки. Эти критерии включают генерацию случайной последовательности, сокрытие распределения, ослепление участников или персонала, затенение оценки результатов и полноту данных о результатах. Риск каждого домена был оценен как высокий, низкий или неясный. Риск для ошибки отбора оценивался как низкий, если при генерации последовательности распределения присутствовал случайный компонент. Риск для сокрытия распределения был оценен как низкий, когда был явно указан четкий метод, который не позволял предвидеть распределение групп до или во время регистрации. Риск для ослепления участников оценивался как низкий, когда исследование включало условие сравнения, которое не позволяло участникам узнать, были ли они отнесены к экспериментальной или контрольной группе (например, приложение-плацебо или вмешательство, предназначенное для терапевтического действия). Риск для ослепления результата оценивали как низкий, если были приняты надлежащие меры для сокрытия принадлежности участников к группе или если результаты были сообщены самостоятельно, что не предполагает прямого контакта с исследователем. Риск для полноты данных о

результатах оценивался как низкий, если авторы исследования включали в свой анализ всех рандомизированных участников (т.е. они придерживались принципа «назначено лечение»).

Мы также извлекли несколько характеристик, касающихся исследования (год, автор, размер выборки), участников (целевая группа, критерии отбора), вмешательства приложения (ориентация, основная цель, ключевые характеристики, предписание человеческого руководства), сравнения (тип контрольного состояния) и оценку результатов (использованный инструмент, оцененная продолжительность, первичный или вторичный). Два исследователя проводили извлечение данных, и любые разногласия разрешались путем консенсуса.

Метаанализ

Анализы были проведены с использованием Комплексного метаанализа версии 3.0[29] (Comprehensive Meta-Analysis Version 3.0) для оценки величины эффекта средних различий между группами и однофакторного анализа подгрупп, а также R для сравнения вариабельности между группами вмешательства и контрольной группой и для изучения взаимодействий между модераторами. Для каждого сравнения средних значений между вмешательством приложения и контрольным состоянием величина эффекта рассчитывалась путем деления разницы между средними значениями двух групп после тестирования на объединенное стандартное отклонение.

Мы указали значение g Хеджеса по сравнению с d Коэна, чтобы скорректировать небольшое отклонение выборки[30]. Если средние значения и стандартные отклонения не сообщались, размеры эффекта рассчитывались с использованием оценок изменений или другой представленной статистики (значения t или *p* для групповых сравнений). Чтобы рассчитать совокупную величину эффекта, размер эффекта каждого исследования был взвешен по его обратной дисперсии. Если использовалось несколько показателей данной переменной результата, вычислялось среднее значение величины эффекта для каждого показателя в рамках исследования, прежде чем величины эффекта были объединены[29]. Положительный g указывает на то, что состояние приложения достигло более высокого уровня редукции симптомов по сравнению с условием сравнения. Величина эффекта 0,8 была интерпретирована как большая, в то время как величина эффекта 0,5 – как умеренная, а величина эффекта 0,2 – как малая[31]. В протоколе мы указали, что проведем метаанализ показателей ремиссии, выздоровления и достоверных изменений. Однако, поскольку эти результаты были редкими (<10% подходящих исследований) и сообщались непоследовательно, мы не смогли проанализировать эти данные.

Был проведен метаанализ различий в вариабельности пост-тестирования для приложений и контрольных групп, поскольку он дает представление о том, являются ли эффекты вмешательства достаточно однородными. Если оценки вариабельности для группы приложения сопоставимы с оценками, полученными для контрольной группы, или меньше, чем они были получены для контрольной группы, предполагается, что вмешательство может иметь хороший потенциал для обобщения. Напротив, большая вариабельность для группы приложений предполагает, что эффекты могут быть ограничены подмножеством участников[24]. Мы провели это сравнение оценок вариабельности, получив логарифмически преобразованную оценку отношения дисперсии группы приложений к дисперсии контрольной группы в результатах пост-тестирования[24]. Альтернативные способы количественной оценки различий в вариабельности[24] были проверены на надежность наших первоначальных результатов. Значимость и направление различий в дисперсии между группами оставались неизменными независимо от используемой операционализации. Мы также провели несколько других анализов чувствительности для оценки надежности вышеуказанных основных результатов. Мы пересчитали совокупные эффекты, ограничив анализ: a) более низким риском систематической ошибки в исследованиях (определяется как соответствие 4 или 5 критериям качества); б) более крупными выборочными исследованиями (определяется как 75 или более рандомизированных участников для каждого условия); в) исследованиями, предоставляющими приложение, которое было специально разработано для устранения симптомов депрессии или тревоги, или когда депрессия или тревога были объявлены в качестве основного исхода; d) наименьший и наибольший эффект в каждом исследовании, если использовалось несколько условий (для поддержания статистической независимости); и e) различная продолжительность посттестового периода (1–4 недели, 5–12 недель или 13 и более недель). Мы также объединили эффекты, исключив выбросы и используя подход неперекрывающихся доверительных интервалов (ДИ), в котором исследование определяется как выброс, когда 95% ДИ величины эффекта не перекрывается с 95% ДИ размера совокупного эффекта[32]. Систематическая ошибка небольших исследований также изучалась с помощью метода обрезки и заполнения[33].

Оценки величины эффекта между группами были дополнены оценками числа больных, которых необходимо лечить (ЧБНЛ), чтобы выразить практическое влияние средневзвешенного значения эффектов вмешательства. ЧБНЛ указывает количество дополнительных участников в группе вмешательства, которым необходимо будет пройти лечение, чтобы наблюдать за одним участником, у которого определяется положительная динамика симптомов по сравнению с контрольной группой[34].

Мы также рассчитали средневзвешенный показатель выбывания из условий приложений включенных испытаний. Он был определен как доля участников, использовавших приложение, которые не прошли посттестовую оценку, деленная на общее количество участников, рандомизированных в соответствии с этим условием. Частота событий, рассчитанная с помощью Комплексного метаанализа (Comprehensive Meta-Analysis), была преобразована в проценты для удобства чтения.

Поскольку ожидалась значительная гетерогенность исследований, для всех анализов использовались модели случайных эффектов[29]. Неоднородность исследовали путем расчета статистики I^2, которая количественно определяет гетерогенность, выявленную статистикой Q, и сообщает, какая общая дисперсия (0–100%) связана с дисперсией между исследованиями[35]. Мы провели серию однофакторных анализов подгрупп, изучая эффекты вмешательства в соответствии с основными характеристиками участников, функциями приложения и исследований (см. также дополнительную информацию). Анализ подгрупп проводился по модели смешанного воздействия[29].

Наконец, признавая, что детерминанты величины эффекта могут взаимодействовать сложным образом, что не может быть адекватно отражено с помощью методов однофакторного анализа подгрупп, мы оценили интерактивное влияние модераторов на оценки эффекта с помощью meta-CART[26]. При этом используется список потенциальных модераторов и предпринимается попытка разделить баллы по ключевой переменной результата (в данном случае – величины эффекта от каждого испытания) в соответствии с комбинациями этих модераторов, которые максимизируют различия между группами в результате, одновременно минимизируя внутригрупповую дисперсию. Этот процесс разделения продолжается до тех пор, пока набор величин эффекта не сможет быть значительно улучшен за счет дальнейшего разделения на подгруппы. Мы использовали 10-кратную перекрестную проверку и моделирование случайных эффектов, учитывая, что ожидается наличие в анализе нескольких источников изменчивости на величину эффекта. Прошлые исследования[36] показывают, что meta-CART хорошо справляется с обнаружением взаимодействия между потенциальными модераторами в случаях, когда в выборке имеется не менее 80 оценок – условие, соблюденное в настоящем обзоре.

Рисунок 1. Блок-схема PRISMA. РКИ – рандомизированное контролируемое исследование

РЕЗУЛЬТАТЫ

Характеристики исследования

На Рисунке 1 представлена блок-схема поиска литературы. В общей сложности 176 РКИ из 174 статей полностью соответствовали критериям включения (см. дополнительную информацию о характеристиках отдельных исследований). Более двух третей приемлемых исследований (67%) были проведены в период с 2020 по 2023 год. Во многих исследованиях (43%) использовалась неотобранная случайная выборка, в то время как в меньшее число исследований были включены люди с депрессией или генерализованной тревожностью, которые либо соответствовали диагностическим критериям (6%), либо набрали баллы выше порогового значения по валидированному показателю (26%). В меньшем количестве исследований принимали участие люди с посттравматическим стрессом, социальной тревожностью, обсессивно-компульсивными или паническими симптомами (10%). Почти половина представленных приложений (48%) была основана на принципах когнитивно-поведенческой терапии (КПТ); меньшее количество приложений было основано на «осознанности» (mindfulness) (21%) или когнитивном тренинге (10%). Треть приложений (34%) имели функции мониторинга настроения, в то время как только 5% включали технологию чат-ботов. В 14% представленных приложений предлагалось руководство со стороны человека.

В большинстве испытаний был предоставлен неактивный контроль (60%), состоящий из списка ожидания, только оценки или информационных ресурсов. В меньшем количестве исследований (23%) использовался плацебо-контроль (например, нетерапевтическое приложение, мгновенные экологичные оценки), в котором предпринималась попытка контролировать время, внимание или ожидания участников. В 11% исследований была оказана обычная медицинская помощь. Только в десяти исследованиях использовалось активное психологическое сравнение, такое как очные лечебные сессии или веб-программа. В большинстве исследований использовалось краткосрочное наблюдение в течение 1–4 (56%) или 5–12 (40%) недель.

Риск системной ошибки также варьировался. Во всех испытаниях использовался показатель депрессии или тревоги по самоотчету; 70% соответствовали критериям для адекватной генерации последовательности; 25% соответствовали критериям для адекватного сокрытия распределения; 26% соответствовали критериям достаточного ослепления участников для условий исследования, и 62% сообщили об использовании анализа намерения лечить. Немногие исследования (6%) соответствовали всем пяти критериям, 21% соответствовали четырем критериям, 32% соответствовали трем критериям, 31% соответствовали двум критериям и 8% соответствовали только одному критерию.

Влияние на симптомы депрессии

Приложения в сравнении с контрольными условиями

Суммарная величина эффекта для 181 сравнения приложений (N=16,569) и контрольных условий (N=17,007) на симптомы депрессии составила g=0,28 (95% ДИ: 0,23–0,33, p<0,001; I^2=72%, 95% ДИ: 67–75), что соответствует ЧБНЛ 11,5 (см. Таблицу 1). Объединенная оценка отношения дисперсии (RoV) составила -0,14 (95% ДИ: от -0,24 до -0,05), что указывает на меньшую дисперсию посттестовых результатов для группы вмешательства по приложению в сравнении с контрольной группой. Однако гетерогенность была высокой для различий в дисперсии (I^2=78%), что указывает на изменчивые эффекты вмешательств приложений.

Анализ чувствительности подтвердил основные выводы (см. Таблицу 1). Совокупная величина эффекта была аналогичной при ограничении анализа меньшим риском систематической ошибки и исследованиями с более крупной выборкой; при поправке на погрешность небольшого исследования в соответствии с процедурой обрезки и заполнения; при ограничении исследованиями, в которых депрессия была основной целью или результатом вмешательства; и при разной продолжительности наблюдения. При исключении выбросов совокупная величина эффекта также была аналогичной, и неоднородность существенно уменьшилась. В ходе анализа чувствительности гетерогенность была высокой для различий в дисперсии, что также указывает на изменчивые эффекты вмешательств приложений (см. дополнительную информацию).

Таблица 1. Метаанализ влияния приложений на симптомы депрессии и генерализованной тревоги

Анализ	Симптомы депрессии	Симптомы генерализованной тревоги								
	n	g (95% ДИ)	*p*	I^2	ЧБНЛ	n	g (95% ДИ)	*p*	I^2	ЧБНЛ
Приложения против контрольных условий	181	0.28 (0.23–0.33)	<0.001	72%	11.5	150	0.26 (0.21–0.31)	<0.001	64%	12.4
Только исследования с низким риском системной ошибки	48	0.32 (0.23–0.40)	<0.001	80%	9.9	35	0.25 (0.18–0.33)	<0.001	49%	12.9
Исследования с небольшими выборками удалены	60	0.22 (0.15–0.28)	<0.001	79%	14.9	47	0.22 (0.16–0.28)	<0.001	67%	14.9
Основная цель или результат вмешательства	60	0.38 (0.28–0.48)	<0.001	84%	8.1	45	0.20 (0.13–0.28)	<0.001	53%	16.5
Выбросы удалены	147	0.25 (0.21–0.28)	<0.001	24%	12.9	131	0.23 (0.20–0.27)	<0.001	16%	14.2
Один эффект на исследование (наименьший)	145	0.27 (0.22–0.32)	<0.001	73%	11.9	121	0.25 (0.20–0.31)	<0.001	69%	12.9
Один эффект на исследование (наибольший)	146	0.31 (0.25–0.35)	<0.001	75%	10.2	121	0.29 (0.24–0.35)	<0.001	68%	11.0
Процедура обрезки и заполнения	177	0.29 (0.24–0.33)	–	75%	11.0	148	0.26 (0.21–0.31)	–	64%	12.4
Продолжительность наблюдения										
1–4 недели	102	0.23 (0.17–0.29)	<0.001	65%	14.2	95	0.21 (0.16–0.26)	<0.001	53%	15.6
5–12 недель	76	0.35 (0.27–0.43)	<0.001	70%	8.9	54	0.34 (0.24–0.44)	<0.001	74%	9.2
≥13 недель	3	0.29 (-0.17 до 0.76)	0.214	87%	11.0	1	0.29 (-0.13 до 0.73)	0.180	0%	11.0
Приложения против активного вмешательства	8	-0.08 (-0.25 до 0.08)	0.340	0%	–	6	0.11 (-0.24 до 0.47)	0.537	64%	31.0
Только очный компаратор	5	-0.12 (-0.35 до 0.09)	0.257	0%	–	5	0.16 (-0.25 до 0.59)	0.441	65%	20.9
Только веб-компаратор	3	-0.01 (-0.31 до 0.29)	0.962	16%	–	1	-0.11 (-0.52 до 0.29)	0.575	0%	-

n – число сравнений, ЧБЛН – число больных, которых необходимо лечить.

В Таблице 2 представлены результаты однофакторного анализа подгрупп. Величина эффекта при депрессии была значительно больше, когда в исследованиях использовалась неактивная контрольная группа (по сравнению с плацебо или обычным лечением), изучалась предварительно отобранная выборка (по сравнению с неотобранной выборкой), применялось приложение для КПТ (по сравнению с приложением, не использующим КПТ) и предоставлялось приложение, содержащее технологию чат-бота (относительно отсутствия технологии чат-ботов).

Мы повторно рассчитали однофакторный анализ подгрупп только среди исследований, в которых депрессия была основной целью вмешательства или исходом (см. Таблицу 3). В ходе этих предварительных анализов были выявлены те же эффекты для подгрупп, за одним исключением: испытания, в которых использовалось приложение для «осознанности» (mindfulness), дали значительно меньший эффект, чем те, в которых приложение для «осознанности» (mindfulness) не использовалось.

Анализ meta-CART выявил пять ключевых особенностей исследования (предварительно отобранная выборка, размер выборки, контрольная группа, психиатрический диагноз и предоставление приложения для когнитивного тренинга), которые характеризовали подгруппы с более высокими или более низкими оценками эффекта для средних различий, чем для выборки в целом. Выборка оценок величины эффекта для депрессии сначала была разделена в зависимости от того, была ли выборка предварительно отобрана. Оценки эффекта от выборок, которые не были предварительно отобраны (g=0,22, 95% ДИ: 0,18–0,27, N=136), не могли быть разделены на дополнительные подгруппы. Величины эффектов для исследований с участием предварительно отобранных выборок были дополнительно разделены на подгруппы, отражающие: а) исследования с более крупной выборкой (≥75 для каждого состояния) плюс контрольные группы плацебо (g=0,06, 95% ДИ: от -0,14 до 0,25, N=6); б) исследования с более крупной выборкой плюс неактивные контрольные группы (g=0,54, 95% ДИ: 0,35–0,74, N=6); в) выборочные исследования меньшего размера плюс выборки с психиатрическим диагнозом (g=0,14, 95% ДИ: от -0,19 до 0,47, N=4); d) выборочные испытания меньшего размера плюс выборки без психиатрического диагноза плюс приложения, которые не были ориентированы на когнитивное обучение (g=0,59, 95% ДИ: от -0,46 до 0,72, N=26); е) выборочные испытания меньшего размера плюс выборки без психиатрического диагноза плюс приложение для когнитивного обучения (g=1,16, 95% ДИ: 0,73–1,60, N=3) (см. также дополнительную информацию).

Приложения в сравнении с активным вмешательством

Совокупная величина эффекта для восьми сравнений между приложениями и активными вмешательствами составила g = -0,08 (95% ДИ: от -0,25 до 0,08, *p*=0,340). Величина эффекта составила g = -0,12 (95% ДИ: от -0,35 до 0,09, *p*=0,257) при ограничении анализа сравнениями с очными лечением и g=-0,01 (95% ДИ: от -0,31 до 0,29, *p*=0,962) при ограничении анализа сравнениями с вмешательством при помощи веб-сайтов, хотя количество исследований в этих анализах было невелико (см. Таблицу 1).

Влияние на симптомы генерализованной тревожности

Приложения в сравнении с контрольными условиями

Совокупная величина эффекта для 150 сравнений между приложениями (N=10,972) и контрольными условиями (N=11,422) на симптомы генерализованной тревожности составила g=0,26 (95% ДИ: 0,21–0,31, *p*<0,001; I^2=64%, 95% ДИ: 57–69), что соответствует значению ЧБНЛ, равному 12,4 (см. Таблицу 1). Суммарная оценка RoV состави-

Таблица 2. Анализ подгрупп по всем доступным исследованиям

Анализ	Симптомы депрессии					Симптомы генерализованной тревоги				
	n	g (95% ДИ)	I^2	ЧБНЛ	*p*	n	g (95% ДИ)	I^2	ЧБНЛ	*p*
Контрольная группа					0.003					0.216
Неактивные	112	0.33 (0.27–0.39)	68%	9.5		96	0.28 (0.23–0.33)	48%	11.4	
Плацебо	49	0.19 (0.11–0.27)	70%	17.4		38	0.19 (0.11–0.28)	55%	17.4	
Обычное лечение	20	0.18 (0.07–0.27)	49%	18.4		16	0.32 (0.09–0.54)	88%	9.2	
Выборка					<0.001					0.080
Предварительно отобранная	45	0.52 (0.38–0.65)	88%	5.2		32	0.35 (0.23–0.46)	62%	8.9	
Неотобранная	136	0.21 (0.17–0.26)	56%	15.6		118	0.24 (0.18–0.29)	64%	13.5	
Психиатрический диагноз					0.669					0.068
Да	14	0.23 (0.05–0.40)	60%	14.2		10	0.41 (0.24–0.58)	35%	7.5	
Нет	167	0.28 (0.23–0.33)	73%	11.4		140	0.25 (0.20–0.30)	65%	12.9	
Приложение КПТ					0.003					0.147
Да	86	0.35 (0.28–0.42)	79%	8.9		59	0.30 (0.24–0.36)	46%	10.6	
Нет	95	0.21 (0.15–0.27)	59%	15.6		91	0.23 (0.16–0.30)	69%	14.2	
Приложение «осознанности» (mindfulness)					0.549					0.258
Да	43	0.26 (0.19–0.33)	45%	12.4		45	0.23 (0.15–0.30)	46%	14.2	
Нет	138	0.29 (0.23–0.34)	75%	11.0		105	0.28 (0.22–0.35)	68%	11.4	
Приложение для когнитивного тренинга					0.238					0.370
Да	21	0.18 (0.02–0.35)	66%	18.4		18	0.20 (0.07–0.33)	39%	16.5	
Нет	160	0.29 (0.24–0.34)	73%	11.0		132	0.26 (0.21–0.31)	65%	12.4	
Функция мониторинга настроения					0.257					0.369
Да	65	0.24 (0.19–0.30)	44%	13.5		51	0.23 (0.17–0.29)	33%	14.2	
Нет/Не сообщалось	116	0.29 (0.23–0.36)	78%	11.0		99	0.27 (0.20–0.34)	71%	11.9	
Функция чат-бота					0.009					0.258
Да	12	0.53 (0.33–0.74)	61%	5.6		10	0.18 (0.06–0.31)	0%	18.4	
Нет/Не сообщалось	169	0.26 (0.21–0.31)	72%	12.4		140	0.26 (0.21–0.32)	66%	12.4	
Предложено руководство человеком					0.936					0.477
Да	27	0.29 (0.13–0.46)	65%	11.0		18	0.37 (0.03–0.71)	88%	8.4	
Нет/Не сообщалось	154	0.28 (0.23–0.32)	73%	11.4		132	0.24 (0.20–0.29)	52%	13.5	

n – число сравнений, ЧБЛН – число больных, которых необходимо лечить, КПТ – когнитивно-поведенческая терапия.

ла -0,21 (95% ДИ: от -0,31 до -0,12), что указывает на меньшую дисперсию посттестовых результатов в группе приложений по сравнению с контрольной группой. Однако гетерогенность была высокой для различий в дисперсии (I^2=75%), что свидетельствует об изменчивой эффективности вмешательств приложений. Эффекты были аналогичными при проведении обширного анализа чувствительности, представленного в Таблице 1.

В однофакторном анализе подгрупп всех доступных исследований, в которых сообщалось о генерализованной тревожности в качестве исхода, не было обнаружено значимых модерирующих факторов (см. Таблицу 2). Однако при ограничении этих анализов исследованиями, где основной целью была генерализованная тревожность, выявилось несколько однофакторных модерирующих эффектов: исследования, в которых использовался неактивный контроль (по сравнению с плацебо или обычным лечением), предварительно отобранные участники по симптомам генерализованной тревожности (по сравнению с неотобранной выборкой), применялось приложение КПТ (по сравнению с приложениями, не связанными с КПТ), а также внедрение приложения с функциями мониторинга настроения привело к значительно большему влиянию на симптомы генерализованной тревожности.

Напротив, исследования, в которых использовалось приложение для тренировки «осознанности» (mindfulness) или когнитивного тренинга, показали значительно меньший эффект на общие симптомы тревоги (см. Таблицу 3).

Анализ meta-CART выявил три ключевых фактора (предварительный отбор, была ли генерализованная тревожность основной целью/исходом, и приложения с функциями мониторинга настроения), которые характеризовали подгруппы с более высокими или более низкими оценками эффекта средних различий, чем для выборки в целом. Выборка оценок величины эффекта сначала была разделена в зависимости от того, была ли выборка предварительно отобрана. Для выборок, которые были предварительно отобраны, оценки эффекта были дополнительно разделены в зависимости от того, включало ли приложение функции мониторинга настроения (g=0,54, 95% ДИ: 0,36–0,27, N=12) или нет (g=0,26, 95% ДИ: 0,13–0,39, N=20).

Для выборок, которые не были предварительно отобраны, оценки эффекта были разделены в зависимости от того, была ли тревога основной целью/результатом (g=0,08, 95% ДИ: от -0,05 до 0,19, N=21) или нет (g=0,28, 95% ДИ: 0,22–0,34, N=97) (см. также дополнительную информацию).

Таблица 3. ***Post-hoc*** **анализ подгрупп по исследованиям, в которых депрессия или генерализованная тревожность были основной целью или исходом вмешательства**

Анализ	Симптомы депрессии					Симптомы генерализованной тревоги				
	n	g (95% ДИ)	I^2	ЧБНЛ	*p*	n	g (95% ДИ)	I^2	ЧБНЛ	*p*
Контрольная группа					0.005					0.050
Неактивные	29	0.56 (0.37–0.74)	82%	5.2		25	0.30 (0.18–0.41)	47%	10.6	
Плацебо	21	0.30 (0.16–0.44)	82%	10.6		15	0.12 (0.02–0.21)	32%	28.3	
Обычное лечение	10	0.16 (0.01–0.31)	67%	20.9		5	0.11 (-0.10 до 0.33)	72%	31.0	
Выборка					<0.001					<0.001
Предварительно отобранная	40	0.55 (0.40–0.70)	88%	5.4		24	0.34 (0.22–0.47)	61%	9.2	
Неотобранная	20	0.12 (0.04–0.21)	35%	28.3		21	0.09 (0.03–0.15)	0%	38.2	
Психиатрический диагноз					0.364					0.071
Да	4	0.18 (-0.26 до 0.63)	67%	18.4		7	0.40 (0.16–0.64)	46%	7.7	
Нет	56	0.40 (0.29–0.50)	85	7.7		38	0.17 (0.09–0.24)	46%	19.6	
Приложение КПТ					0.041					0.029
Да	40	0.46 (0.32–0.58)	87%	6.6		22	0.28 (0.18–0.38)	41%	11.4	
Нет	20	0.24 (0.09–0.40)	78%	13.5		23	0.12 (0.01–0.23)	52%	28.3	
Приложение «осознанности» (mindfulness)					0.024					0.036
Да	9	0.21 (0.07–0.35)	31%	15.6		9	0.07 (-0.06 до 0.21)	24%	49.5	
Нет	51	0.42 (0.31–0.53)	85%	7.3		36	0.24 (0.15–0.33)	56%	13.5	
Приложение для когнитивного тренинга					0.218					0.016
Да	4	0.85 (0.07–1.63)	83%	3.3		7	0.03 (-0.10 до 0.17)	52%	117.5	
Нет	56	0.35 (0.26–0.45)	83%	8.9		38	0.23 (0.15–0.31)	19%	14.2	
Функция мониторинга настроения					0.327					0.033
Да	26	0.33 (0.21–0.45)	65%	9.5		25	0.28 (0.18–0.37)	44%	11.4	
Нет/Не сообщалось	34	0.42 (0.27–0.57)	88%	7.3		20	0.12 (0.01–0.22)	52%	28.3	
Функция чат-бота					0.005					0.493
Да	5	0.80 (0.50–1.10)	47%	3.5		5	0.21 (0.13–0.29)	0%	15.6	
Нет/Не сообщалось	55	0.34 (0.25–0.44)	84%	9.2		40	0.14 (-0.04 до 0.33)	56%	24.1	
Предложено руководство человеком					0.503					0.358
Да	8	0.52 (0.11–0.92)	80%	5.7		3	0.45 (-0.10 до 1.01)	84%	6.7	
Нет/Не сообщалось	52	0.37 (0.27–0.47)	85%	8.4		42	0.19 (0.11–0.26)	47%	17.4	

n – число сравнений, ЧБЛН – число больных, которых необходимо лечить, КПТ – когнитивно-поведенческая терапия.

Приложения в сравнении с активным вмешательством

Совокупная величина эффекта для шести сравнений между приложениями и активными вмешательствами составила g=0,11 (95% ДИ: от -0,24 до 0,47, *p*=0,537). Величина эффекта составила g=0,16 (95% ДИ: от -0,25 до 0,59, *p*=0,441) при ограничении анализа сравнениями с очным лечением и g=-0,11 (95% ДИ: от -0,52 до 0,29, *p*=0,575) при ограничении анализа одним сравнением с веб-вмешательство (см. Таблицу 1).

Влияние на специфические симптомы тревоги

Влияние приложений по сравнению с контрольными условиями на характерные симптомы тревоги представлено в Таблице 4. Анализ подгрупп и сравнение приложений с активными вмешательствами по этим результатам не проводились из-за ограниченного числа доступных исследований.

Совокупная величина эффекта для 17 сравнений между приложениями (N=1,371) и контрольными условиями (N=1,385) при симптомах посттравматического стресса составила g=0,12 (95% ДИ: 0,03–0,21, *p*=0,007), что соответствует ЧБНЛ 28,3. Гетерогенность была низкой (I^2=24%). Значимые небольшие эффекты наблюдались во всех анализах чувствительности, кроме случаев, когда были исключены исследования с меньшими выборками.

Совокупная величина эффекта для десяти сравнений между приложениями (N=576) и контрольными условиями (N=447) на симптомы социальной тревожности составила g=0,52 (95% ДИ: 0,22–0,82, *p*=0,001), что соответствует ЧБНЛ 5,7. Гетерогенность была высокой (I^2=75%). Эффекты оставались стабильными и сходными по величине, когда анализ ограничивался испытаниями, в которых предварительно отбирались участники по симптомам социальной тревожности, применялось приложение КПТ и где социальная тревожность была основной целью или исходом. Незначительные эффекты наблюдались при ограничении анализа испытаниями с более низким риском системной ошибки и большей выборкой. Однако только одно исследование было оценено как имеющее низкий риск и более крупную выборку[37].

Таблица 4. Метаанализ влияния приложений на специфичные симптомы тревожности

Исходы	Анализ	n	g (95% ДИ)	*p*	I^2	ЧБНЛ
Симптомы посттравматического стресса	Приложения в сравнении с контрольными условиями	17	0.12 (0.03–0.21)	0.007	24%	28.3
	Только исследования с низким риском системной ошибки	3	0.34 (0.11–0.57)	0.004	22%	9.2
	Исследования с небольшими выборками удалены	5	0.11 (-0.04 до 0.25)	0.150	59%	31.0
	Один эффект на исследование (наименьший)	15	0.13 (0.03–0.23)	0.008	29%	26.0
	Один эффект на исследование (наибольший)	15	0.14 (0.04–0.24)	0.005	30%	24.1
	Процедура обрезки и заполнения	14	0.13 (0.02–0.24)	–	–	26.0
	Посттравматический стресс, основная цель/результат	12	0.12 (0.01–0.24)	0.039	44%	28.3
	Предварительный отбор по симптомам посттравматического стресса	10	0.14 (0.04–0.25)	0.006	29%	24.1
	Только приложения КПТ	12	0.15 (0.02–0.29)	0.019	35%	22.4
Симптомы социальной тревожности	Приложения в сравнении с контрольными условиями	10	0.52 (0.22–0.82)	0.001	75%	5.7
	Только исследования с низким риском системной ошибки	1	0.10 (-0.16 до 0.37)	0.446	0%	34.3
	Исследования с небольшими выборками удалены	1	0.10 (-0.16 до 0.37)	0.446	0%	34.3
	Один эффект на исследование (наименьший)	9	0.53 (0.19–0.86)	0.002	77%	5.6
	Один эффект на исследование (наибольший)	9	0.61 (0.28–0.93)	<0.001	71%	4.8
	Процедура обрезки и заполнения	6	0.24 (-0.06 до 0.55)	–	–	13.5
	Посттравматический стресс, основная цель/результат	6	0.74 (0.24–1.24)	0.003	83%	3.9
	Предварительный отбор по симптомам посттравматического стресса	5	0.75 (0.16–1.32)	0.011	86%	3.8
	Только приложения КПТ	4	0.73 (0.01–1.45)	0.044	82%	3.9
Обсессивно-компульсивные симптомы	Приложения в сравнении с контрольными условиями	5	0.51 (0.18–0.84)	0.002	42%	5.8
Симптомы паники	Приложения в сравнении с контрольными условиями	2	-0.12 (-0.50 до 0.25)	0.515	0%	–
Симптомы акрофобии	Приложения в сравнении с контрольными условиями	2	0.90 (0.38–1.42)	0.001	0%	3.0

n – число сравнений, ЧБЛН – число больных, которых необходимо лечить, КПТ – когнитивно-поведенческая терапия.

Значительная совокупная величина эффекта наблюдалась для пяти сравнений между приложениями и контрольными условиями по обсессивно-компульсивным симптомам (g=0,51, 95% ДИ: 0,18–0,84, *p*=0,002; ЧБНЛ=5,8) и для двух сравнений по симптомам акрофобии (g=0,90, 95% ДИ: 0,38–1,42, *p*=0,001; ЧБНЛ=3,0). При двух сравнениях между приложениями и контрольными условиями наблюдался незначительный отрицательный эффект на симптомы паники (g =-0,12, 95% ДИ: от -0,50 до 0,25, *p*=0,515) (см. Таблицу 4).

Частота выбывания

Из 182 состояний, по которым имелись доступные данные, взвешенный показатель отсева составил, по оценкам, 23,6% (95% ДИ: 21,3–26,1, I^2=93%). При исключении исследований с небольшими выборками частота отсева составила 29,9% (95% ДИ: 26,0–34,0, I^2=96%) из 67 случаев. При ограничении анализа исследованиями с низким риском системной ошибки частота отсева составила 26,6% (95% ДИ: 22,7–31,0, I^2=94%) из 50 состояний. Для выборок участников с депрессией частота отсева составила 28,7% (95% ДИ: 23,6–34,3, I^2=93%) при 42 состояниях. Для выборок участников с тревогой частота отсева составила 25,4% (95% ДИ: 20,6–31,0, I^2=92%) из 39 состояний. Для приложений КПТ показатель отсева составил 23,3% (95% ДИ: 19,8–27,2, I^2=94%) из 89 условий.

ОБСУЖДЕНИЕ

Интерес к приложениям для психического здоровья как к масштабируемому инструменту лечения симптомов депрессии и тревоги продолжает расти. Со времени последнего всестороннего метаанализа, опубликованного в 2019 году[13], было проведено более 100 РКИ. Чтобы убедиться, что клиницисты, регулирующие органы здравоохранения и общественность имеют доступ к самой последней информации о доказательной базе этих приложений, мы провели обновленный метаанализ 176 научных исследований. Определенное внимание было уделено выявлению особенностей, которые могут объяснять очевидную и значительную неоднородность эффективности от исследования к исследованию. Это первый метаанализ приложений для психического здоровья, в котором проводится тщательный анализ того, как могут взаимодействовать комбинации предполагаемых факторов, с целью по-новому взглянуть на обстоятельства и подгруппы людей, для которых определенные функции приложения могут иметь наибольший эффект.

В общей сложности, результаты показали, что приложения для психического здоровья оказывают в целом небольшое, но значительное влияние на симптомы депрессии (N=33,576, g=0,28) и генерализованной тревожности (N=22,394, g=0,26), что соответствует ЧБНЛ 11,5 и 12,4 соответственно. Гетерогенность была высокой в основных тестах, но существенно снижалась при удалении выбросов.

При проведении обширного анализа чувствительности эффекты были устойчивыми и сходными по величине при разной продолжительности наблюдения, а также после удаления исследований с небольшими выборками и более высоким риском системных ошибок. Более выраженные эффекты были обнаружены для депрессии, когда она была основной целью приложения (g=0,38), в то время как для генерализованной тревоги подобного не наблюдалось (g=0,20). Отсев был очевидным, и каждый четвертый участник преждевременно выбыл из выделенной для него программы приложений. Ограниченные несущественные эффекты в отношении депрессии и генерализованной тревожности наблюдались при оценке приложений в сравнении с веб и очными вмешательствами, хотя количество исследований было небольшим, а доверительные интервалы широкими.

При пост-тестировании в приложении наблюдалась меньшая вариабельность результатов по сравнению с контрольными условиями (RoV=-0,14 для депрессивных симптомов и RoV= -0,21 для симптомов генерализованной тревоги). Однако гетерогенность была высокой для различий в дисперсии (I^2=78% для депрессивных симптомов и I^2=75% для симптомов генерализованной тревожности), что свидетельствует об изменчивой эффективности вмешательств приложений.

Растущее количество литературы теперь позволяет нам оценить влияние приложений для психического здоровья на специфичные симптомы тревоги и подчеркивает потенциал более специализированных подходов. Наши предыдущие метааналитические оценки влияния приложений на симптомы социальной тревожности, паники и посттравматического стресса были основаны только на трех – шести сравнениях, выявивших ограниченные доказательства эффективности[13]. В настоящее время литературные данные обновляются, и число исследований, нацеленных на определенные симптомы, увеличилось более чем втрое (например, посттравматический стресс), в то же время появились новые исследования, позволяющие рассчитать предварительные совокупные эффекты и для других симптомов (например, обсессивно-компульсивные симптомы, акрофобия). Мы обнаружили доказательства умеренного влияния приложений на симптомы социальной тревожности (N=10, g=0,52) и обсессивно-компульсивные симптомы (N=5, g=0,51), небольшого влияния на симптомы посттравматического стресса (N=17, g=0,12), большого влияния на симптомы акрофобии (N=2, g=0,90) и незначительного негативного влияния на симптомы паники (N=2, g=-0,12). Однако к этим результатам следует относиться с осторожностью, поскольку большинство исследований, способствовавших проведению этих анализов, имели значительный риск системной ошибки и были основаны на небольших размерах выборки.

Наши результаты также показывают, каким образом достижения в методологии проведения исследований помогут лучше оценить эффективность существующих приложений и разработать новые. На однофакторном уровне тип контрольного состояния выступал в качестве модератора, при этом неактивные контрольные группы оказывали более выраженное влияние на депрессию и генерализованную тревожность (определены в качестве основной цели), чем контрольные группы плацебо или обычное лечение. Это легко воспроизводимое открытие, наблюдаемое при всех способах психологического лечения[38-41], оно служит подтверждением того, что некоторые преимущества приложений объясняются эффектом «цифрового плацебо»[42]. Теперь, когда этот эффект плацебо хорошо известен, возникает настоятельная необходимость обеспечить понимание его реальных последствий как врачами, которые оценивают эффективность или «назначают» определенные приложения, так и регулирующими органами, которые сертифицируют их утверждения. Кроме того, эффекты были более выраженными среди предварительно отобранных выборок участников с симптомами депрессии и генерализованной тревожности порогового уровня (идентифицированных в качестве основной цели). Этот вывод согласуется с предыдущими метааналитическими данными о том, что более тяжелое исходное состояние связано с лучшими результатами цифровых вмешательств[43,44], предполагая, что люди с умеренной или тяжелой депрессией или тревогой на исходном уровне получают больше пользы от помощи, дополненной этими приложениями[45].

Особенности приложения также были связаны с величиной эффекта на однофакторном уровне. Приложения, основанные на КПТ, давали больший эффект, чем другие приложения, например, основанные на «осознанности» (mindfulness). Этот вывод неудивителен, учитывая, что доказательная база для всех форм КПТ при депрессии и тревоге значительно больше, чем для других подходов[46]. Возможно, лучшее понимание «активных составляющих» КПТ при этих состояниях (в сравнении с другими подходами) способствовало разработке приложений, которые отдают предпочтение этим эффективным компонентам перед другими, менее эффективными или потенциально вредными.

Кроме того, мы обнаружили некоторые доказательства того, что эффект был сильнее, когда приложения, специально разработанные для борьбы с депрессией, включали технологию чат-ботов, а приложения, специально разработанные для борьбы с тревогой, включали функции мониторинга настроения. Вполне возможно, что эти инновационные технологические функции обеспечивают большую степень персонализации, являются более привлекательными, способствуют эмоциональному самосознанию и повышают ответственность пользователей за достижение прогресса[47,48], что потенциально приводит к большему преимуществу. Однако эти данные были получены в результате *post-hoc* анализа, и их следует рассматривать с осторожностью, поскольку количество исследований в этих подгруппах было относительно небольшим. Необходимы рандомизированные эксперименты, которые проверяют дополнительные эффекты чатбота и технологии мониторинга настроения, механизмы действия и факторы вовлеченности.

На многофакторном уровне комбинации предложенных модерирующих переменных помогли лучше идентифицировать подгруппы, в которых приложения были более или менее эффективны. Что касается депрессии, исследования с более высокими оценками эффективности, как правило, характеризовались предварительно отобранными выборками меньшего размера и без формального психиатрического диагноза. Приложения для когнитивного тренинга показали более высокую эффективность при депрессии в сочетании с этими особенностями исследования (g=1,16), но мы подчеркиваем необходимость осторожности при интерпретации этого эффекта, учитывая малочисленность исследований в этой группе (N=3). При генерализованной тревожности приложения с мониторингом настроения были особенно эффективны в исследованиях с предварительно отобранными выборками, в то время как положительные преимущества этого мониторинга не проявились для более универсальных выборок.

Существуют некоторые ограничения этого метаанализа, которые следует учитывать. Во-первых, анализ проводился только в период после вмешательства из-за непоследовательности отчетов и продолжительности последующих оценок, а также различий в том, как решались проблемы отсева. Таким образом, остается открытым вопрос о том, распространяются ли преимущества приложений, наблюдаемые в краткосрочной перспективе, на более длительные периоды. Во-вторых, анализ можно проводить только по динамике симптомов, а не по другим клинически значимым исходам, таким как ремиссия, выздоровление или ухудшение состояния[49]. Несмотря на достаточное количество исследований, включавших людей с депрессией или тревожностью лишь малая часть исследователей сообщала об этих исходах, а если сообщала, то определялись они непоследовательно. В-третьих, гетерогенность была высокой во многих основных тестах. Анализ с выделением подгрупп не смог объяснить ее полностью, и многие подгруппы попрежнему отличались высокой гетерогенностью.

Однако, возможно, сохранение некоторой неоднородности при агрегировании данных из исследований существенно разных приложений будет всегда неизбежным – независимо от того, насколько хорошо классифицированы отдельные компоненты и характеристики исследований.

В заключение мы представляем наиболее полный мета-анализ влияния приложений для психического здоровья на симптомы депрессии и тревоги, включающий 176 исследований. Мы пришли к выводу, что приложения в целом оказывают небольшое, но значительное влияние на симптомы депрессии и генерализованной тревожности. Более серьезные эффекты наблюдаются в исследованиях, в которых депрессия является основной целью или исходом вмешательства, это позволяет предположить, что приложения могут стать подходящим первым шагом в лечении для лиц, восприимчивых к этому подходу, или для тех, кто не может получить доступ к традиционным формам медицинской помощи. Определенные функции приложений, такие как мониторинг настроения и технология чат-ботов, были связаны с большей величиной эффекта, хотя это необходимо подтвердить в будущих экспериментальных исследованиях. Доказательства, подтверждающие эффективность приложений при специфических симптомах тревоги, являются неопределенными, в основном из-за испытаний со значительным риском системной ошибки и небольшим размером выборки. Поскольку чувствительность к приложениям для психического здоровья различна, сбор и объединение большого объема данных (включая пассивные данные и данные о самоотчетах) будут полезны в дальнейших исследованиях для создания прогностических моделей, способных точно отличать лиц, для которых приложения достаточно, от тех, кто нуждается в других формах лечения.

БЛАГОДАРНОСТИ

J. Linardon удерживает грант National Health and Medical Research Council Grant (APP1196948); J. Torous поддерживается фондом Argosy Foundation; J. Firth поддерживается UK Research and Innovation Future Leaders Fellowship (MR/T021780/1). Дополнительная информация об этом исследовании доступна на сайте: https://osf.io/ufb2w/.

Перевод: Занкишиева Карина Азноровна, г. Санкт-Петербург
Редактура: Камаева Дарья Андреевна, г. Томск

Список литературы

1. Santomauro DF, Herrera AMM, Shadid J et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021;398:1700-12.
2. Liu Q, He H, Yang J et al. Changes in the global burden of depression from 1990 to 2017: findings from the Global Burden of Disease study. J Psychiatr Res 2020;126:134-40.
3. Cuijpers P, Donker T, Weissman MM et al. Interpersonal psychotherapy for mental health problems: a comprehensive meta-analysis. Am J Psychiatry 2016;173:680-7.
4. Cuijpers P, Miguel C, Harrer M et al. Cognitive behavior therapy vs. control conditions, other psychotherapies, pharmacotherapies and combined treatment for depression: a comprehensive meta-analysis including 409 trials with 52,702 patients. World Psychiatry 2023;22:105-15.
5. Carl E, Witcraft SM, Kauffman BY et al. Psychological and pharmacological treatments for generalized anxiety disorder (GAD): a meta-analysis of randomized controlled trials. Cogn Behav Ther 2020;49:1-21.
6. Cipriani A, Furukawa TA, Salanti G et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet 2018;391:1357-66.
7. Cuijpers P, Sijbrandij M, Koole SL et al. The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: a meta-analysis of direct comparisons. World Psychiatry 2013;12:137-48.
8. Torous J, Bucci S, Bell IH et al. The growing field of digital psychiatry: current evidence and the future of apps, social media, chatbots, and virtual reality. World Psychiatry 2021;20:318-35.
9. Torous J, Myrick KJ, Rauseo-Ricupero N et al. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Mental Health 2020;7:e18848.
10. Andersson G. Internet-delivered psychological treatments. Annu Rev Clin Psychol 2016;12:157-79.
11. Hedman-Lagerlöf E, Carlbring P, Svärdman F et al. Therapist-supported Internet-based cognitive behaviour therapy yields similar effects as face-to-face therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. World Psychiatry 2023;22:305-14.
12. Poushter J. Smartphone ownership and internet usage continues to climb in emerging economies. Pew Research Center 2016;22:1-44.
13. Linardon J, Cuijpers P, Carlbring P et al. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019;18:325-36.
14. Goldberg SB, Lam SU, Simonsson O et al. Mobile phone-based interventions for mental health: a systematic meta-review of 14 meta-analyses of randomized controlled trials. PLoS Digit Health 2022;1:e0000002.
15. Baumel A, Torous J, Edan S et al. There is a non-evidence-based app for that: a systematic review and mixed methods analysis of depression-and anxiety-related apps that incorporate unrecognized techniques. J Affect Disord 2020;273:410-21.
16. Larsen ME, Huckvale K, Nicholas J et al. Using science to sell apps: evaluation of mental health app store quality claims. NPJ Digit Med 2019;2:18.
17. Torous J, Myrick K, Aguilera A. The need for a new generation of digital mental health tools to support more accessible, effective and equitable care. World Psychiatry 2023;22:1-2.
18. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphonedelivered interventions for mental health problems: a systematic and metanalytic review. J Consult Clin Psychol 2020;88:1-13.
19. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR mHealth uHealth 2015;3:e4924.
20. Camacho E, Cohen A, Torous J. Assessment of mental health services available through smartphone apps. JAMA Netw Open 2022;5:e2248784.
21. Firth J, Torous J, Nicholas J et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017;16:287-98.
22. Cuijpers P, Griffin JW, Furukawa TA. The lack of statistical power of subgroup analyses in meta-analyses: a cautionary note. Epidemiol Psychiatr Sci 2021;30:e78.
23. Firth J, Torous J, Nicholas J et al. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017;218:15-22.
24. Mills HL, Higgins JPT, Morris RW et al. Detecting heterogeneity of intervention effects using analysis and meta-analysis of differences in variance between trial arms. Epidemiology 2021;32:846-54.
25. Usui T, Macleod MR, McCann SK et al. Meta-analysis of variation suggests that embracing variability improves both replicability and generalizability in preclinical research. PLoS Biology 2021;19:e3001009.
26. Li X, Dusseldorp E, Meulman JJ. Meta-CART: a tool to identify interactions between moderators in meta-analysis. Br J Math Stat 2017;70:118-36.
27. Moher D, Liberati A, Tetzlaff J et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009; 151:264-9.
28. Higgins J, Green S. Cochrane handbook for systematic reviews of interventions. Chichester: Wiley, 2011.
29. Borenstein M, Hedges LV, Higgins JP et al. Introduction to meta-analysis. Chichester: Wiley, 2009.
30. Hedges LV, Olkin I. Statistical methods for meta-analysis. San Diego: Academic Press, 1985.
31. Cohen J. A power primer. Psychol Bull 1992;112:155-9.
32. Harrer M, Cuijpers P, Furukawa TA et al. Doing meta-analysis with R: a hand-son guide. New York: Chapman and Hall/CRC, 2021.
33. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000; 56: 455-63.
34. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. BMJ 1995;310:452-4.
35. Higgins J, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.

36. Li X, Dusseldorp E, Meulman JJ. A flexible approach to identify interaction effects between moderators in meta-analysis. Res Synth Methods 2019;10:134-52.
37. Bruehlman-Senecal E, Hook CJ, Pfeifer JH et al. Smartphone app to address loneliness among college students: pilot randomized controlled trial. JMIR Ment Health 2020;7:e21496.
38. Linardon J. Can acceptance, mindfulness, and self-compassion be learnt by smartphone apps? A systematic and meta-analytic review of randomized controlled trials. Behav Ther 2020;51:646-58.
39. Michopoulos I, Furukawa TA, Noma H et al. Different control conditions can produce different effect estimates in psychotherapy trials for depression. J Clin Epidemiol 2021;132:59-70.
40. Christ C, Schouten MJ, Blankers M et al. Internet and computer-based cognitive behavioral therapy for anxiety and depression in adolescents and young adults: systematic review and meta-analysis. J Med Internet Res 2020; 22: e17831.
41. Linardon J, Shatte A, Messer M et al. E-mental health interventions for the treatment and prevention of eating disorders: an updated systematic review and meta-analysis. J Consult Clin Psychol 2020;88:994-1007.
42. Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. Lancet Psychiatry 2016;3:100-2.
43. 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.
44. Scholten W, Seldenrijk A, Hoogendoorn A et al. Baseline severity as a moderator of the waiting list-controlled association of cognitive behavioral therapy with symptom change in social anxiety disorder: a systematic review and individual patient data meta-analysis. JAMA Psychiatry 2023;80:822-31.
45. van Straten A, Hill J, Richards D et al. Stepped care treatment delivery for depression: a systematic review and meta-analysis. Psychol Med 2015;45:231-46.
46. David D, Cristea I, Hofmann SG. Why cognitive behavioral therapy is the current gold standard of psychotherapy. Front Psychiatry 2018;9:4.
47. Torous J, Lipschitz J, Ng M et al. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. J Affect Disord 2020;263:413-9.
48. He Y, Yang L, Qian C et al. Conversational agent interventions for mental health problems: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2023;25:e43862.
49. Cuijpers P, Karyotaki E, Ciharova M et al. The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: a meta-analysis. Acta Psychiatr Scand 2021;144:288-99.

DOI:10.1002/wps.21183