Разработка и временная валидация клинической прогностической модели развития психоза у лиц с ультравысоким риском в когорте UHR 1000+

Simon Hartmann[1-3], Dominic Dwyer[2,3], Blake Cavve[4], Enda M. Byrne[5], Isabelle Scott[2,3], Caroline Gao[2,3], Cassandra Wannan[2,3], Hok Pan Yuen[2,3], Jessica Hartmann[2,3,6], Ashleigh Lin[5], Stephen J. Wood[2,3,7], Johanna T.W. Wigman[8], Christel M. Middeldorp[9-13], Andrew Thompson[2,3], Paul Amminger[2,3], Monika Schlögelhofer[14], Anita Riecher-Rössler[15], Eric Y.H. Chen[16-18], Ian B. Hickie[19], Lisa J. Phillips[20], Miriam R. Schäfer[2,3], Nilufar Mossaheb[21], Stefan Smesny[22], Gregor Berger[23], Lieuwe de Haan[24], Merete Nordentoft[25,26], Swapna Verma[18,27], Dorien H. Nieman[23], Patrick D. McGorry[2,3], Alison R. Yung[28], Scott R. Clark[1], Barnaby Nelson[2,3]

[1]Discipline of Psychiatry, Adelaide Medical School, University of Adelaide, Adelaide, SA, Australia; [2]Orygen, Melbourne, VIC, Australia; [3]Centre for Youth Mental Health, University of Melbourne, Melbourne, VIC, Australia; [4]Telethon Kids Institute, University of Western Australia, Perth, WA, Australia; [5]School of Population and Global Health, University of Western Australia, Perth, WA, Australia; [6]Department of Public Mental Health, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany; [7]School of Psychology, University of Birmingham, Birmingham, UK; [8]Department of Psychiatry, Interdisciplinary Center Psychopathology and Emotion Regulation, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; [9]Child Health Research Centre, University of Queensland, St. Lucia, QLD, Australia; [10]Department of Child and Adolescent Psychiatry and Psychology, Amsterdam University Medical Center, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands; [11]Arkin Mental Health Care, Amsterdam, The Netherlands; [12]Levvel, Academic Center for Child and Adolescent Psychiatry, Amsterdam, The Netherlands; [13]Child and Youth Mental Health Service, Children's Health Queensland Hospital and Health Service, Brisbane, QLD, Australia; [14]BioPsyC – Biopsychosocial Corporation, Vienna, Austria; [15]Medical Faculty, University of Basel, Basel, Switzerland; [16]Department of Psychiatry, School of Clinical Medicine, LKS Faculty of Medicine, University of Hong Kong, Hong Kong, Hong Kong; [17]LKC School of Medicine, Nanyang Technological University, Singapore, Singapore; [18]Institute of Mental Health, Singapore, Singapore; [19]Brain and Mind Centre, University of Sydney, Sydney, NSW, Australia; [20]Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, VIC, Australia; [21]Department of Psychiatry and Psychotherapy, Clinical Division of Social Psychiatry, Medical University of Vienna, Vienna, Austria; [22]Department of Psychiatry, Jena University Hospital, Jena, Germany; [23]Department of Child and Adolescent Psychiatry and Psychotherapy, Psychiatric Hospital, University of Zurich, Zurich, Switzerland; [24]Department of Psychiatry, Academic Medical Center, Amsterdam, The Netherlands; [25]Mental Health Center Copenhagen, Research Unit (CORE), Copenhagen, Denmark; [26]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; [27]Office of Education, Duke-NUS Graduate Medical School, Singapore, Singapore; [28]Institute for Mental and Physical Health and Clinical Translation, Deakin University, Geelong, VIC, Australia.

Концепция ультравысокого риска (UHR) развития психоза находится в центре внимания исследований в области психиатрии на протяжении нескольких десятилетий. Ее целью является предотвращение развития психотического расстройства у лиц с высоким риском. Исследовательский центр Orygen (Мельбурн, Австралия) провел ряд обсервационных и интервенционных исследований в данной клинической популяции. Эти наборы данных были объединены в когорту UHR 1000+, включающую 1245 человек с UHR и периодом наблюдения от 1 до 16,7 года. Данная статья описывает эту когорту, представляет модель клинического прогнозирования развития психоза в данной группе и анализирует, как изменения в выборках UHR со временем влияют на точность прогноза. Мы проанализировали развитие психоза с использованием модели пропорциональных рисков Кокса. Клинические предикторы развития психоза исследовались во всей когорте с использованием метода множественной импутации и правил Рубина. Для оценки изменения точности модели с течением времени данные за 1995–2016 годы использовались для первоначальной настройки модели, а данные за 2017–2020 годы – для ее последующей валидации. В течение периода наблюдения у 220 человек (17,7%) развилось психотическое расстройство. Объединенные показатели отношения рисков (HR) выявили следующие предикторы развития психоза: тяжесть подшкалы «Дезорганизованная речь» по шкале CAARMS (HR=1.12, 95% ДИ: 1.02-1.24, p=0.024); тяжесть подшкалы «Необычное содержание мыслей» по шкале CAARMS (HR=1.13, 95% ДИ: 1.03-1.24, p=0.009); общий балл по шкале оценки негативных симптомов (SANS) (HR=1.02, 95% ДИ: 1.00-1.03, p=0.022); балл по шкале оценки социального и профессионального функционирования (SOFAS) (HR=0.98, 95% ДИ: 0.97-1.00, p=0.036); время от начала симптомов до обращения в службу UHR (логарифмически преобразованное) (HR=1.10, 95% ДИ: 1.02-1.19, p=0.013). У участников с кратковременные интермиттирующими психотическими симптомами (BLIPS) вероятность развития психоза была выше, чем у тех, кто соответствовал критериям аттенуированных психотических симптомов (APS) (HR=0.48, 95% ДИ: 0.32-0.73, p=0.001), и у тех, кто соответствовал критериям генетического и конституционального риска (наличие близкого родственника с психотическим расстройством или наличие у пациента шизотипического расстройства личности и значительное снижение функциональности в течение прошлого года) (HR=0.43, 95% ДИ: 0.22-0.83, p=0.013). Модели, основанные на данных 1995–2016 годов, продемонстрировали хорошую калибровку при начальной настройке, но отклонение на 20.2–35.4% при проверке на данных за 2017–2020 годы. Для разработки точных моделей прогнозирования психоза необходимы данные масштабных лонгитюдных исследований, таких как UHR 1000+. Важно оценивать существующие и будущие калькуляторы риска на предмет временного отклонения (temporal drift), которое со временем может снижать ценность моделей в клинической практике.

Ключевые слова: ультравысокий риск психоза, прогнозирование психоза, UHR 1000+, дезорганизованная речь, необычное содержание мыслей, негативные симптомы, психосоциальное функционирование, кратковременные интермиттирующие психотические симптомы, калибровка прогностической модели, временная валидация

Simon Hartmann, Dominic Dwyer, Blake Cavve et al. Development and temporal validation of a clinical prediction model of transition to psychosis in individuals at ultra-high risk in the UHR 1000+ cohort. *World Psychiatry*. 2024;23(3):400-410

Критерии ультравысокого риска (ultra high risk, UHR) развития психоза были разработаны в Австралии в середине 1990-х годов для выявления людей с высоким риском развития психотического расстройства и создания превентивных стратегий лечения[1-7]. С тех пор критерии UHR, также известные как критерии «клинического высокого риска» (clinical high risk, CHR), получили широкое признание во всем мире и в настоящее время являются наиболее распространенным методом идентификации людей с высоким риском развития психотического расстройства[8].

Люди, относящиеся к группе UHR, определяются по наличию одного или нескольких из следующих признаков: а) ослабленные психотические симптомы (attenuated psychotic symptoms, APS) – то есть наличие в течение последнего года подпороговых, ослабленных форм продуктивных психотических симптомов; б) кратковременные интермиттирующие психотические симптомы (brief limited intermittent psychotic symptoms, BLIPS) – то есть наличие эпизодов развернутых психотических симптомов, продолжающихся не более недели и спонтанно разрешающихся; в) конституцио-

нальный и генетический риск (Trait) – то есть наличие близкого родственника с психотическим расстройством или наличие у пациента шизотипического расстройства, а также значительное снижение уровня функционирования за последний год.

Согласно недавнему метаанализу, 25% людей, соответствующих критериям UHR, переживают первый психотический эпизод в течение трехлетнего периода[9], что свидетельствует о значительно более высоком риске по сравнению с общей популяцией и другими клиническими группами (0,05% и 3,9% соответственно[10]). Исходы для пациентов UHR остаются гетерогенными, отсюда необходимость более точного расчета риска на индивидуальном уровне, для чего могут потребоваться мультимодальные данные[11].

Было проведено множество исследований, направленных на совершенствование методов выявления людей с UHR, улучшение их прогноза и разработку индивидуализированных стратегий вмешательства[12]. Согласно метаанализу, прогностические модели развития психотических расстройств у лиц с UHR достигли чувствительности 67% и специфичности 78%[13]. В последние годы совместные усилия национальных и международных мультицентровых проектов, таких как PRONIA (www.pronia.eu), PSYSCAN (www.psyscan.eu) и NAPLS-3 (http://campuspress.yale.edu/napls), привели к созданию масштабных баз данных для дальнейшего развития точных прогностических моделей. Однако большинство этих когорт имеют короткий период наблюдения (1–2 года), что затрудняет оценку долгосрочных клинических исходов и изучение кратковременных изменений в выборках UHR. Таким образом, существует необходимость в масштабных базах данных UHR, содержащих мультимодальные данные и охватывающих длительные периоды наблюдения.

Для получения более убедительных доказательств для их последующего внедрения акцент сместился в сторону оценки транспортабельности и возможности генерализовать предиктивные модели, что включает внешнюю валидацию предиктивной точности на выборках из разных географических регионов и популяций пациентов, а также применение различных методов измерения предикторов[17]. Однако точность прогностических моделей развития психоза зависит не только от региональных различий, но и от изменений в выборках ультравысокого риска с течением времени. Эти изменения связаны с эволюцией диагностических и терапевтических моделей, изменением маршрутизации пациентов и трансформацией состава популяций, что может проявляться в снижении частоты развития психоза в группе риска[18-21]. В результате ранее валидированные прогностические модели могут терять актуальность, так как основаны на данных, отражавших ситуацию на тот момент времени, и могут не соответствовать изменениям в популяции, клинических условиях и подходах к лечению[22].

Например, временная валидация клинической модели прогнозирования смертности после кардиоторакальной хирургии (модели «EuroSCORE»), выявила снижение ее точности с течением времени, что выражалось в расхождении между наблюдаемой и прогнозируемой смертностью[23]. На индивидуальном уровне риск смертности был завышен при использовании модели, основанной на более старых данных, что потенциально могло привести к чрезмерно консервативному подходу в принятии хирургических решений. Таким образом, временная валидация необходима для оценки стабильности моделей на протяжении разных временных интервалов (месяцев, лет и, в конечном итоге, десятилетий) в одних и тех же клинических условиях.

Мало известно о том, как изменения в популяциях UHR и методах лечения с течением времени влияют на эффективность моделей прогнозирования. В одном из предыдущих исследований[24] была представлена мультимодальная модель прогнозирования психоза на основе машинного обучения, для которой посередине периода сбора данных выборка была разделена на две части. Этот подход напоминает попытку проведения временной валидации, но изучение отклонений в модели с течением времени не было целью исследования.

Еще одним ограничением многих предыдущих моделей является акцент на прогностической точности без оценки точности индивидуальных показателей риска. Модели могут выявлять у пациента определенный уровень риска, но прогнозируемый риск может отличаться от наблюдаемого риска, то есть модель может быть плохо откалибрована. Лишь небольшое количество исследований проанализировало калибровку моделей прогнозирования психоза[25-29]. Плохая калибровка может влиять на целесообразность решений, принимаемых с использованием модели, когда выбор вмешательства основывается на пороговой вероятности (то есть на стратификации людей в группы по прогнозируемому риску). Хотя такие процессы принятия решений обычно включают предпочтения пациента и отношение врача к выявленному риску[30], их можно представить более упрощенно с помощью кривых чистой выгоды (то есть баланса между риском и выгодой вмешательства при заданной пороговой вероятности)[26,29,31]. Определение потенциального изменения точности и особенностей калибровки с течением времени критически важно для обоснования внедрения прогностических моделей развития психоза в клиническую практику.

Целью данного исследования является изучение точности прогнозирования психоза с использованием большого набора данных и оценка стабильности индивидуального прогноза риска развития психоза на протяжении определенного времени. Мы объединили данные более чем 1000 человек с ультравысоким риском, собранные в рамках исследований, проведенных в исследовательском центре Orygen (Мельбурн) в период с 1995 по 2020 год, что сформировало крупнейшую на сегодняшний день UHR-когорту с самым продолжительным периодом лонгитюдного наблюдения.

МЕТОДЫ

Выборка

Выборка включала лиц с ультравысоким риском, участвовавших в исследованиях центра Orygen (Мельбурн) с 1995 по 2020 год. Исследования в основном проводились в Мельбурне на базе клиники Персонального обследования и кризисной оценки (Personal Assessment and Crisis Evaluation) (PACE)[32], Центра профилактики и раннего вмешательства при психозе (Early Psychosis Prevention and Intervention Centre) (EPPIC) и четырех центров Headspace (Sunshine, Craigieburn, Werribee и Glenroy). Одно из исследований также собирало дополнительные данные (198 человек) на площадках в Амстердаме, Базеле, Копенгагене, Йене, Гонконге, Мельбурне, Сингапуре, Сиднее, Вене и Цюрихе[33]. В таблице 1 представлены 14 исследований, в рамках которых были набраны участники данной когорты (пять клинических исследований[33-38] и девять продольных обсервационных исследований[3, 39-46]) (см. дополнительную информацию).

Критерии включения и исключения

Критерии включения: а) способность дать информированное согласие; б) соответствие критериям хотя бы одной из трех групп UHR (см. дополнительную информацию). Критерии исключения для всех исследований: а) текущее или ранее перенесенное психотическое расстройство или маниакальный эпизод; б) прием антипсихотиков в дозе, эквивалентной более 15–50 мг галоперидола (в зависимости от оригинального исследования) в анамнезе; в) психотическое расстройство, вызванное употреблением психоактивных веществ в анамнезе; г) наличие медицинских состояний, которые могли бы объяснить симптомы, послужившие причиной направления (например, эпилепсия); д) диагноз выраженного расстройства нейропсихологического развития; е) задокументированная задержка развития или умственная отсталость (IQ<70); ж) недостаточное владение английским языком для участия в исследовании.

Статус ультравысокого риска (UHR)

В период с 1995 по 1996 год статус UHR на исходном уровне оценивался с помощью Шкалы краткой психиатри-

Таблица 1. Исследования, в которых были набраны лица, включенные в когорту UHR 1000+

Исследование	Годы набора	Возраст в исходной точке	UHR (N)	Временные точки
Yung et al[3]	1995–1996	14–30 лет	49	BL, 1–11 месяцы
McGorry et al[34]	1996–1999	14–30 лет	92	BL, 6–12 месяцы
Thompson et al[39]	2000	14–30 лет	39	BL, 1–11 месяцы
Berger et al[36]	2000–2005	14–30 лет	30	BL, 1–23 месяцы
Yung et al[35]	2000–2007	14–30 лет	111	BL, 1–23 месяцы
Phillips et al[40]	2000–2005	14–30 лет	78	BL, 1–23 месяцы
Nelson et al[46]	2008–2010	15–25 лет	49	BL
Amminger et al[33] McGorry et al[38]	2010–2014	13–40 лет	304	BL, 1–12–24 месяцы
EU-GEI[43]	2012–2015	15–35 лет	35	BL, 12–24 месяцы
Nelson et al[45]	2014–2018	15–25 лет	48	BL, 12 месяц
Hartmann et al[42]	2016–2018	12–25 лет	31	BL, 12 месяц
Nelson et al[37]	2016–2019	12–25 лет	342	BL, 1–3–6–9–12–18–24 месяцы
Tognin et al[44]	2017–2019	15–25 лет	26	BL, 3–6–9–12–18 месяцы
Bayer et al[41]	2018–2021	12–25 years	12	BL, месяц 12

UHR – состояние ультравысокого риска, BL – исходный уровень.

ческой оценки (Brief Psychiatric Rating Scale, BPRS)[47], Комплексной оценки симптомов и истории болезни (Comprehensive Assessment of Symptoms and History, CASH)[48] и Глобальной шкалы оценки функционирования (Global Assessment of Functioning, GAF)[49]. С 1996 по 1999 год применялись Шкала комплексной оценки состояний повышенного риска развития психических расстройств (Comprehensive Assessment of At-Risk Mental States, CAARMS)[50] и Шкала глобального функционирования (Global Assessment of Functioning, GAF). Начиная с 1999 года статус UHR устанавливался с использованием CAARMS и Шкалы оценки социального и профессионального функционирования (Social and Occupational Functioning Assessment Scales, SOFAS)[51].

Показатели исхода

Основным интересующим исходом было развитие психотического расстройства, определяемое с использованием соответствующих критериев на момент исследования. До 1999 года психотический порог устанавливался на основании данных BPRS/CASH и CAARMS, пока не была установлена совокупная достоверность CAARMS. С 1999 года BPRS/CASH как основа для определения психоза были заменены на CAARMS. Для случаев, в которых данные CAARMS были недоступны, для выявления факта развития психоза использовались государственные записи об оказании психиатрической помощи[52].

Возможные предикторы развития психоза

Развитие психотического расстройства анализировалось с помощью модели пропорциональных рисков Кокса. Переменные были выбраны на основе предыдущих исследований и включали: показатели тяжести по шкале продуктивных симптомов CAARMS (дезорганизованная речь, нарушения восприятия и необычное содержание мыслей); продуктивные симптомы, количественно оцененные по общему баллу BPRS; негативные симптомы, оцененные по общему баллу шкалы оценки негативных симптомов (SANS)[53]; уровень функционирования, оцененный по шкале GAF или SOFAS; длительность симптомов до обращения в службу UHR; пол; возраст в исходной точке; подгруппа UHR (BLIPS, APS или Trait) (см. дополнительную информацию).

Мы также включили переменную, указывающую, получал ли участник стандартное или нестандартное лечение в рамках клинического исследования, чтобы выявить потенциальные различия, связанные с лечением. Стандартное лечение включало ведение случая, симптоматическое лечение и коррекцию дистресса. Примеры нестандартных вмешательств включали лечение ω-3 полиненасыщенными жирными кислотами или когнитивно-поведенческую терапию (КПТ) и рисперидон.

Индивиды с UHR были стратифицированы по степени тяжести симптомов, используя группы включения CAARMS: BLIPS, APS и Trait. Лица, соответствующие более чем одной категории, классифицировались по самой высокой степени тяжести: любой BLIPS; APS или APS + Trait; только Trait. Показатели GAF были преобразованы в оценки SOFAS с использованием эквипроцентильного метода[54]. Длительность симптомов до обращения в службу UHR была логарифмически преобразована перед анализом.

Статистический анализ

Пропущенные данные обрабатывались с использованием метода множественной импутации (multiple imputation by chained equations, MICE) при условии, что пропущенные данные распределены случайно[55] (см. дополнительную информацию). Модель Кокса была построена для каждого замененного (импутированного) набора данных; объединенные оценки для каждого возможного предиктора рассчитывались с использованием правил Рубина[56,57]. Время наблюдения рассчитывалось от даты начальной оценки до развития психоза или выхода из исследования (пациент пропал из поля зрения или исследование было официально завершено), в зависимости от того, какой случай наступал раньше. Далее была разработана полная модель пропорциональных рисков Кокса, прогнозирующая развитие психоза в течение 2 и 10 лет, используя клинические переменные, которые были идентифицированы как предикторы развития психоза на предыдущем этапе и прошли внутреннюю валидацию с помощью бутстрэппинга (N=1000). Бутстрэппинг включает построение нескольких моделей (в данном случае 1000) с использованием случайных выборок с заменой данных исследования для определения эффективности модели на той же выборке пациентов[58]. Оба временных промежутка (2 и 10 лет) использовались для оценки эффективности модели как в краткосрочной, так и в долгосрочной перспективе.

Когорта была разделена по годам начальной оценки на набор для разработки модели (1995–2016) и набор для ее валидации (2017–2020). Кроме того, был проанализирован сокращенный набор данных, включающий только исследования из выборки PACE 400 (1995–2007)[52], чтобы смоделировать более ранний набор данных. Первоначально модели настраивались на каждом замещенном (импутированном) наборе данных с использованием данных за 1995–2016 и 1995–2007 годы. После чего они проходили внутреннюю

валидацию методом бутстрэппинга (N=1000) для оценки их точности, а затем временную валидацию на данных за 2017–2020 годы. Временная валидация проводилась путем прогнозирования двухлетней вероятности развития психоза среди лиц, обратившихся в службу UHR в Мельбурне, чтобы исключить возможные региональные различия в наборе данных.

Для каждой модели изменение точности оценивалось путем сравнения показателей дискриминационной способности, значений калибровки и анализа кривой принятия решений, полученных в результате внутренней и временной валидации.

Дискриминационная способность – это способность модели присваивать более высокий прогнозируемый риск людям, у которых развился психоз, по сравнению с теми, у кого психоз не развился[59]. Хотя дискриминационная способность важна для определения того, может ли модель различать людей с высоким и низким риском, не менее важным является оценка калибровки прогностической модели (т.е. насколько близки наблюдаемая и прогнозируемая вероятности). Например, модель, которая корректно различает двух пациентов, присваивая второму пациенту риск в два раза выше, чем первому, может быть плохо откалибрована, если прогнозируемые вероятности составляют 40% и 20%, а фактически наблюдаемые вероятности – 10% и 5%[60].

Анализ кривых принятия решений представляет собой метод оценки прогностических моделей[31], в котором соотношение пользы и вреда клинического решения рассматривается в контексте истинно положительного показателя для заданного прогнозируемого порога вероятности модели. Это соотношение рассматривается как «чистая выгода» и сравнивается со сценариями, в которых лечатся все или никто[61]. Данные кривые основаны на фактической вероятности событий и превосходят положительную прогностическую ценность (ППЦ) и отрицательную прогностическую ценность (НПЦ), которые чувствительны к ошибкам, если исходный уровень заболеваемости изменяется со временем[30]. Более высокая чистая выгода означает, что модель идентифицирует больше истинно положительных случаев, которые затем могут получить необходимое лечение. В прогнозировании психоза польза может заключаться в предотвращении развития полноценного психотического эпизода в группе ультравысокого риска, а вред может включать ненужное лечение и стигматизацию[26].

Анализ проводился с использованием R версии 4.1.173 с использованием пакетов mice[62], survival[63,64] и rms[65].

РЕЗУЛЬТАТЫ

Характеристики выборки

В таблице 2 представлены исходные характеристики групп, у которых развился и не развился психоз в когорте UHR 1000+. В общей сложности в исследование было включено 1245 участников UHR. Средний возраст участников составил 18,5 ± 3,6 года; 55% из них – женщины. Среднее время от начала симптомов до обращения в службу UHR составило 646,6 ± 1021,7 дня. Нестандартное лечение в рамках клинического исследования получили 667 участников (54%). Наибольшая группа участников соответствовала критериям APS (85,2%), либо исключительно, либо в сочетании с критериями Trait. За ними следовали участники, соответствующие только критериям Trait (7,8%). Лишь 86 участников соответствовали критериям BLIPS (6,9%).

Частота развития психотического расстройства

Всего у 220 из 1245 участников группы ультравысокого риска развился психоз (17,7%). Частота развития психоза среди участников каждого клинического исследования, включенного в данную когорту, представлена в таблице 3. Медиана времени до развития психоза составила 210 дней (межквартильный размах, IQR = 500,5 дня). Два участника были исключены из анализа, так как дата манифестации психоза была неизвестна.

Возможные предикторы развития психоза

Как показано на рисунке 1, следующие факторы предсказывали развитие психоза в общей когорте UHR 1000+: степень тяжести показателя «Дезорганизованная речь» по шкале CAARMS (отношение рисков, HR=1,12, 95% ДИ: 1,02-1,24, p=0,024); степень тяжести показателя «Необычное содержание мыслей» по шкале CAARMS (HR=1,13, 95% ДИ: 1,03-1,24, p=0,009); общий балл по шкале SANS (HR=1,02, 95% ДИ: 1,00-1,03, p=0,022); балл по шкале SOFAS (HR=0,98, 95% ДИ: 0,97-1,00, p=0,036); время от начала симптомов до обращения в службу UHR (логарифмически преобразованное) (HR=1,10, 95% ДИ: 1,02-1,19, p=0,013). Время от начала симптомов до обращения в службу UHR следует интерпретировать с осторожностью из-за большого объема замещенных (импутированных) данных (~20%).

Участники UHR, соответствующие критериям BLIPS, имели более высокую частоту развития психоза, чем те, кто соответствовал критериям APS или APS + Trait (HR=0,48, 95% ДИ: 0,32-0,73, p=0,001), и те, кто соответствовал только критериям Trait (HR=0,43, 95% ДИ: 0,22-0,83, p=0,013).

Нестандартное лечение, проводимое в рамках клинического исследования, было связано со снижением риска развития психоза по сравнению со стандартным лечением (HR=0,74, 95% ДИ: 0,56-0,98, p=0,038). Однако при добавлении идентификатора исследования, прогностическое значение переменной лечения стало несущественным (HR=1,18, 95% ДИ: 0,80-1,72, p=0,40). Это, скорее всего, связано с тем, что данные из обсервационных исследований были объединены с данными клинических исследований, при этом первые проводились раньше и показатели частоты развития психоза в них были выше. В связи с этим переменная лечения была исключена из дальнейшего анализа.

Внутренняя валидация прогностической модели психоза

Мы построили множественную модель регрессии Кокса, включающую в качестве предикторов: степень тяжести показателя «Дезорганизованная речь» по шкале CAARMS (от 0 до 6); степень тяжести показателя «Необычное содержание мыслей» по шкале CAARMS (от 0 до 6); время между началом симптомов и обращением в службу UHR (логарифмически преобразованное); общий балл по шкале SANS (от 0 до 100); балл по шкале SOFAS (от 0 до 100); группу UHR на момент включения.

На основе заданного набора значений предикторов модель вычисляет непрерывный риск развития психоза в диапазоне от 0 до 1 для каждого индивида (см. дополнительную информацию). Согласно внутренней валидации с использованием бутстрэппинга, модель достигла умеренного уровня дискриминационной способности (точности, C-индекса) в размере 0,65. Это означает, что она успешно определяла более высокий риск примерно в двух третях случаев. Внутренняя валидация выявила низкую степень переобучения модели (6–7%), что свидетельствует о том, что модель сохраняет точность в широком диапазоне клинических проявлений выборки (см. дополнительные данные для кривых калибровки и анализа кривых принятия решений).

«Пороговое значение вероятности» позволяет установить порог для прогнозирования развития психоза в заданный период времени: пациенты с показателем выше порога, согласно прогнозу, перенесут психоз в течение этого периода времени, тогда как те, у кого показатель ниже, – нет. Это позволяет эффективно разделить выборку на две группы риска. Для определения оптимального порога, точность прогнозирования была протестирована на различных пороговых уровнях и оценена через 2 и 10 лет наблюдения. В таблице 4 приведена статистика прогнозирования для разных порогов риска.

При пороговом значении 0,20 (т.е. вероятность развития психоза у индивида составляет 20%) положительная прогностическая ценность (ППЦ) модели (доля людей, точно идентифицированных как те, у которых манифестирует

Таблица 2. Социально-демографические и клинические характеристики лиц, развивших или не развивших психоз, на исходном уровне

	Отсутствие развития психоза (N=1,025)	Развитие психоза (N=220)
Возраст, лет (среднее значение ± СО)	18,4±3,6	18,9±3,7
Женщины (%)	55,7	52,7
С супругом или гражданским партнером (%)	4,2	7,4
С детьми (%)	3,8	5,4
Среднее образование (%)		
Продолжающееся	44,0	45,1
Завершенное	31,1	28,8
Техническое и дополнительное образование	11,0	16,3
Незавершенное обучение по программам бакалавриата	10,5	7,1
Завершенное обучение по программам бакалавриата	2,5	1,1
Незавершенное обучение по программам магистратуры и аспирантуры	0,6	1,1
Завершенное обучение по программам магистратуры и аспирантуры	0,2	0,5
Иммигрировали из другой страны (%)	14,0	8,3
Один или оба родителя иммигрировали из другой страны (%)	39,8	40,6
Время от начала симптомов до обращения в службу UHR, дней (среднее значение ± СО)	621,5±983,4	754,5±1168,8
Подгруппа UHR (%)		
BLIPS	5,5	13,7
APS или APS + Trait	86,5	79,5
Trait	8,0	6,8
Получали нестандартное лечение в рамках исследования	55	45
CAARMS, дезорганизованная речь, степень тяжести (среднее значение ± СО)	1,6±1,4	2,2±1,4
CAARMS, нарушения восприятия, степень тяжести (среднее значение ± СО)	3,3±1,5	3,3±1,7
CAARMS, необычное содержание мыслей, степень тяжести (среднее значение ± СО)	2,8±1,8	3,5±1,6
Общий балл BPRS (среднее значение ± СО)	43,9±9,2	48,2±10,7
Общий балл SANS (среднее значение ± СО)	17,6±11,8	22,9±13,3
Балл SOFAS (среднее значение ± СО)	56,7±11,7	53,3±10,9

UHR – состояние ультравысокого риска, BLIPS – короткие интермиттирующие психотические симптомы, APS – аттенуированные психотические симптомы, CAARMS – комплексная оценка состояний риска развития психических расстройств, BPRS – шкала краткой психиатрической оценки, SANS – шкала оценки негативных симптомов, SOFAS – шкала оценки социального и профессионального функционирования.

психоз) составила 29,72%, при отрицательной прогностической ценности (ОПЦ) – 87,91%, для прогнозирования развития психоза в течение двух лет. Для 10-летнего прогноза ППЦ составила 22,26%, а ОПЦ – 91,56%. При пороговом значении риска 0,50 ППЦ увеличилась до 49,69%, в то время как ОПЦ снизилась до 82,67% для двухлетнего прогноза, а для 10-летнего прогноза ППЦ составила 39,35%, а ОПЦ – 84,52%.

Эти показатели указывают на то, что модель более полезна для исключения риска развития психоза у индивида (высокая ОПЦ), чем для выявления случаев высокого риска (ППЦ). Повышение порога риска увеличивало ППЦ, но снижало ОПЦ. Увеличение периода прогнозирования (10 лет по сравнению с 2 годами) увеличивало ОПЦ, но снижало ППЦ.

Временная валидация прогностической модели развития психоза

Таблица 5 содержит сводные данные о калибровке и дискриминационной способности для внутренней и временной валидации моделей, построенных с использованием тех же предикторов, что и на предыдущем этапе, на выборках с 1995 по 2016 год и с 1995 по 2007 год (PACE 40052). Кривые калибровки и анализ кривых принятия решений для обеих моделей представлены на рисунке 2.

Обе модели показали хорошую калибровку после внутренней валидации в рамках временного диапазона их разработки, с уровнем переобучения около 9% (9,4% для модели PACE 400 и 8,3% для модели 1995–2016 годов). Кроме того, обе модели достигли умеренного уровня дискриминационной способности – 0,672–0,684.

При проверке модели на данных 2017–2020 годов из тех же клинических служб обе модели продемонстрировали отклонение в калибровке, как показано на рисунке 2. В обоих случаях калибровочные кривые в наборе для валидации «отклонялись» от идеальной кривой, где прогнозируемая вероятность соответствует наблюдаемой вероятности. Для обеих моделей потеря калибровки составила 20,2–35,4%, что привело к переоценке вероятности развития психоза у лиц в более поздней выборке.

Например, в выборке 1995–2016 годов 30 из 100 человек, у которых прогнозируемая вероятность развития психоза составляла 0,30, действительно перенесли психоз, что говорит о соответствии прогнозируемого и наблюдаемого риска. В выборке 2017–2020 годов калибровочные кривые показывают, что только 25 из 100 человек с аналогичными симптомами и прогнозируемым риском развития психоза 0,30 действительно перенесли психоз, что указывает на расхождение между прогнозируемым и наблюдаемым риском развития психоза.

Потенциальное влияние отклонений на принятие решений представлено на кривой принятия решений чистой выгоды, которая показывает снижение доли истинно положительных случаев (чистой выгоды) для заданного порога вероятности в модели (см. рисунок 2). Это означает, что мо-

Таблица 3. Частота развития психоза и время от первичной оценки до развития психоза

Исследование	Годы набора	Продолжительность наблюдения (лет)	Развитие первого психотического эпизода	Дней от первичной оценки до развития психоза, медиана (межквартильный размах)
Yung et al[3]	1995–1996	6–15	26/48 (54,2%)	120,0 (295,5)
McGorry et al[34]	1996–1999	6–15	36/92 (39,1%)	234,0 (672,75)
Thompson et al[39]	2000	6–15	12/39 (30,8%)	313,5 (1280,5)
Berger et al[36]	2000–2007	6–15	38/219 (17,3%)	363,5 (789,0)
Yung et al[35]				
Phillips et al[4]				
Nelson et al[46]	2008–2010	1.5	13/49 (26,5%)	46,0 (65,0)
McGorry et al[38]	2010–2014	3.4	40/304 (13,1%)	229,0 (288,75)
EU-GEI[43]	2012–2015	2	6/35 (17,1%)	131,0 (87,75)
Nelson et al[45]	2014–2018	1	3/48 (6,2%)	77,0 (11,0)
Hartmann et al[42]	2016–2018	1	2/31 (6,4%)	31,0 (9,0)
Nelson et al[37]	2016–2019	2	37/342 (10,8%)	220,0 (409,5)
Tognin et al[44]	2017–2019	2	4/26 (15,4%)	257,5 (55,5)
Bayer et al[41]	2018–2021	1	3/12 (25,0%)	173,0 (111,5)
Всего			220/1245 (17,7%)	210,0 (500,5)



Рисунок 1. Возможные предикторы развития психоза (отношение рисков с 95% ДИ). CAARMS – комплексная оценка состояний риска развития психических расстройств, BPRS – шкала краткой психиатрической оценки, SANS – шкала оценки негативных симптомов, SOFAS – шкала оценки социального и профессионального функционирования, UHR – состояние ультравысокого риска, BLIPS – короткие интермиттирующие психотические симптомы, APS – аттенуированные психотические симптомы

дель завышает риск психоза (например, истинная вероятность составляет 0,25, в то время как прогнозируемая – 0,30) в более поздней выборке (2017–2020 годы).

Предполагается, что лица с показателем выше порога риска разовьют психоз в течение указанного периода времени, в то время как те, у кого показатель ниже порога, – нет. ОПЦ – отрицательная прогностическая ценность, ППЦ – положительная прогностическая ценность.

С-индекс Харрелла используется для количественной оценки дискриминационной способности модели. Глобальная калибровка (calibration-in-the-large) отражает разницу между средним прогнозируемым риском и общей частотой событий, а наклон калибровки показывает расхождение между прогнозируемым и наблюдаемым риском. Внутренняя валидация проводилась методом бутстрэппинга (N=1000) на выборках за 1995–2007 и 1995–2016 годы. Временная валидация заключалась в прогнозировании двухлетней вероятности развития психоза у пациентов, обратившихся в службу UHR в период с 2017 по 2020 год, с использованием моделей, разработанных на выборках 1995–2007 и 1995–2016 годов

ОБСУЖДЕНИЕ

Относительно низкий уровень развития психоза среди лиц с высоким риском вызвал повышенный интерес к разработке и валидации клинических моделей прогнозирования, позволяющих более точно оценивать индивидуальный риск. Несмотря на реализацию крупных когортных исследований, таких как PRONIA и NAPLS-3, необходимо про-

Таблица 4. Статистика прогнозирования развития психоза при различных пороговых значениях риска с использованием бутстрэппинга (N=1000)

Порог риска	Прогнозирование развития психоза в течение 2 лет				Прогнозирование развития психоза в течение 10 лет			
	Специфичность (%)	Чувствительность (%)	ОПЦ (%)	ППЦ (%)	Специфичность (%)	Чувствительность (%)	ОПЦ (%)	ППЦ (%)
0,05	3,10	99,54	96,98	17,93	0	100	0	17,54
0,10	28,64	88,00	91,82	20,78	4,86	98,18	92,65	18,00
0,15	53,44	66,89	88,36	23,41	19,01	91,77	91,58	19,42
0,20	73,68	52,32	87,91	29,72	37,92	83,55	91,56	22,26
0,25	86,38	37,63	86,69	37,04	51,60	67,96	88,33	23,00
0,30	91,65	24,34	85,07	38,31	64,81	60,11	88,43	26,66
0,35	95,88	11,93	83,66	38,12	77,97	48,86	87,72	31,98
0,40	97,86	7,30	83,23	42,02	85,49	38,06	86,65	35,83
0,45	99,42	4,97	83,11	64,37	90,56	29,30	85,76	39,80
0,50	99,61	1,82	82,67	49,69	93,65	19,34	84,52	39,35
0,55	99,61	1,36	82,60	42,47	96,08	11,59	83,63	38,67
0,60	99,90	0,90	82,58	65,28	97,87	7,87	83,32	44,02

Таблица 5. Обобщенные показатели точности внутренней и временной валидации для обеих моделей, рассчитанные на всех импутированных наборах данных (все значения представлены в виде средних с 95% ДИ)

Модель	С-индекс Харрела (индекс конкордации)		Глобальная калибровка (calibration-in-the-large)		Наклон	
	Внутренний	2017–2020	Внутренний	2017–2020	Внутренний	2017–2020
1995-2007	0,684 (0,683–0,686)	0,615 (0,611–0,618)	0,010 (0,010–0,010)	-0,075 (-0,079 до -0,071)	0,906 (0,904–0,908)	0,585 (0,565–0,604)
1995-2016	0,672 (0,671–0,674)	0,636 (0,633–0,638)	0,009 (0,008–0,009)	-0,023 (-0,023 до -0,022)	0,917 (0,916–0,919)	0,732 (0,717–0,747)

водить аналогичные исследования с более крупными выборками и длительным периодом наблюдения. Это позволит изучить изменения в популяциях UHR со временем и проследить долгосрочные клинические исходы.

Мы объединили и стандартизировали данные о людях из группы UHR, набранных в 14 клинических исследованиях, создав когорту UHR 1000+, представленную в данном исследовании. Это крупнейший на сегодняшний день массив клинических данных о людях с ультравысоким риском развития психоза. Большая часть данных (84%) была собрана в одном географическом регионе (Мельбурн, Австралия), что существенно снижает влияние различий между местами проведения исследований. Кроме того, более чем у 50% участников этой когорты наблюдение длилось свыше двух лет, что делает ее крупнейшей выборкой UHR с длительным наблюдением.

В рамках всей когорты объединенные баллы по таким показателям, как «Дезорганизованная речь» и «Необычное содержание мыслей» по CAARMS, негативные симптомы (общий балл по SANS), социальное функционирование (оценка по шкале SOFAS) и длительность симптомов до обращения в службу UHR, служили предикторами развития психоза. Мы также выявили более низкий риск развития психоза у лиц, соответствующих критериям APS или Trait, по сравнению с теми, кто соответствовал любым критериям BLIPS. Это подтверждает выводы предыдущего метаанализа[66].

Получение нестандартного лечения в рамках клинического исследования было связано с более низким риском развития психоза во всей когорте, что согласуется с данными метаанализа о положительном влиянии превентивных вмешательств на риск развития психоза в клинических исследованиях[67]. Однако дальнейший анализ показал, что эта ассоциация, по-видимому, была обусловлена объединением данных обсервационных и клинических исследований, имевших разные уровни развития психоза в разные годы. Таким образом, расчеты, объединяющие эти выборки, являются ненадежными; необходим дальнейший, более подробный анализ прогностических моделей в рамках масштабного продольного наблюдения за группами из клинических исследований.

Принятие клинических решений для пациентов с UHR представляет сложность из-за разнородности клинических проявлений и сравнительно низких показателей развития психоза. В практике сложные модели оценки риска могут быть адаптированы в виде упрощенных калькуляторов, позволяющих занятым клиницистам использовать ограниченный набор данных для помощи в принятии индивидуальных решений. Ряд подобных калькуляторов был предложен для оценки риска развития психоза[25-29]. Как правило, в таких инструментах устанавливается числовой порог, служащий сигналом для вмешательства: если рассчитанная вероятность события, основанная на комбинации факторов риска, превышает установленное значение, это сигнализирует о необходимости вмешательства. Данные пороговые значения можно определить с помощью таких метрик, как точность, чувствительность и специфичность, или ППЦ и ОПЦ, которые помогают клиницистам понять, насколько хорошо калькулятор классифицирует пациентов с высоким и низким риском.

Наша модель, разработанная на основе когорты UHR 1000+, показала умеренную точность с дискриминационной способностью 0,65. Модель также продемонстрировала высокую внутреннюю валидность и хорошую калибровку в выборке, на которой она была построена, что свидетельствует о ее надежности и точности в данном контексте. Возможность генерализации этой модели еще предстоит проверить за пределами выборки UHR 1000+. Объединение клинических данных с биомаркерами риска развития психоза, полученными с помощью методов нейровизуализа-

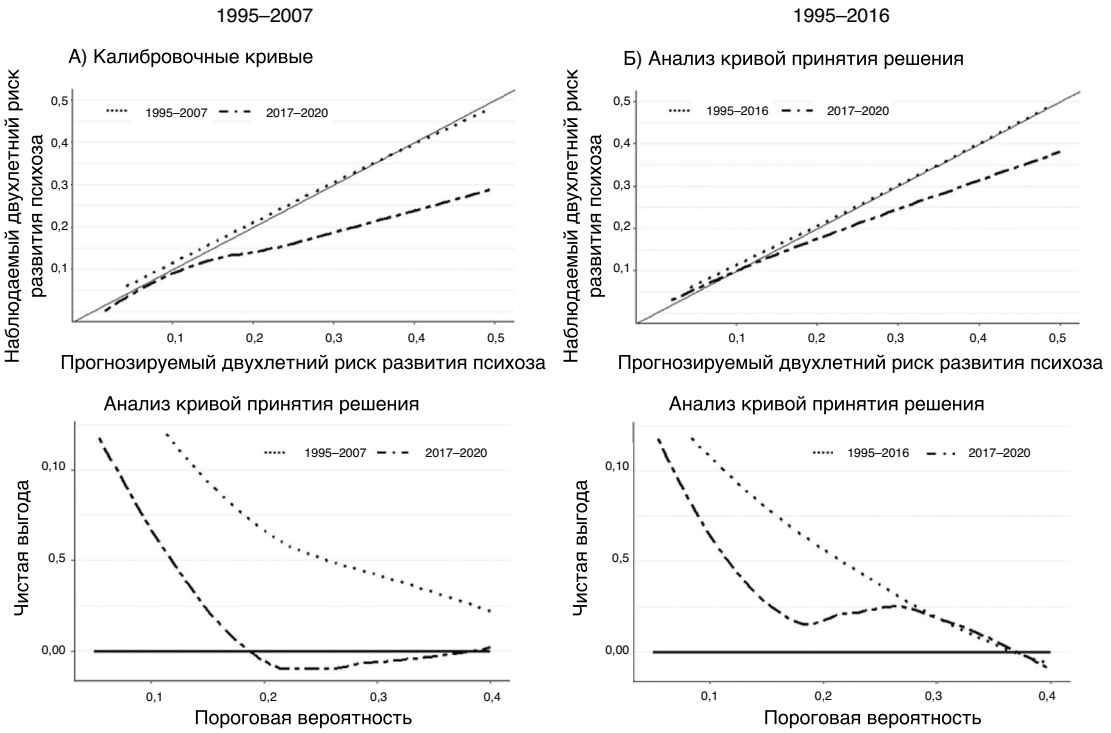


Рисунок 2. Изменение точности прогностических моделей развития психоза с течением времени, разработанных на выборках с 1995 по 2007 и с 1995 по 2016 год. А) Каждый график отображает внутреннюю калибровочную кривую, построенную методом бутстрэппинга (N=1000), и кривую временной валидации, основанную на данных пациентов, обратившихся в службу UHR в период с 2017 по 2020 год. Черная диагональная линия представляет идеальную калибровку, при которой прогнозируемые вероятности точно соответствуют наблюдаемым вероятностям. В) Каждый график отображает анализ кривой принятия решений для внутренней и временной выборки. В этом анализе чистая выгода эквивалентна числу истинно положительных случаев (т.е. чистая выгода 0,10 означает выявление 10 из 100 человек, у которых действительно манифестирует психоз). Черная горизонтальная линия обозначает чистую выгоду, равную нулю, что эквивалентно сценарию, при котором лечение не назначается никому

ции, когнитивной науки, геномики и цифровой оценки, может повысить точность модели[24,68-70], в настоящее время как раз проводятся масштабные исследования, оценивающие этот подход[71]. Кроме того, принятие решений в клинической практике является сложным процессом, который требует учитывать предпочтения пациента и зависит от отношения врача к риску вмешательства[30].

При временной валидации наших моделей на более позднем наборе данных (2017–2020 годы) мы выявили снижение калибровки на 20,2–35,4%, что свидетельствует о переоценке риска психоза, хотя модели показали хорошую калибровку и высокую воспроизводимость на выборке, на которой они были разработаны. Этот результат подчеркивает, что помимо оценки дискриминационной способности калькуляторов риска в разных когортах и клинических службах[72], врачи должны учитывать давность разработки модели прогнозирования развития психоза при интерпретации ее результатов. Эти выводы согласуются с данными из других областей медицины, где калькуляторы риска требуют периодического обновления для поддержания точности калибровки[73-76].

Расхождение между прогнозируемым и наблюдаемым риском развития психоза в когорте UHR 1000+ может быть обусловлено различием частоты развития психоза в разных исследованиях за этот период времени (22% в выборке 1995–2016 годов против 11% в выборке 2017–2020 годов). Это может быть связано с проблемами выборки. Со временем в этой когорте увеличилось количество участников, набранных через headspace – службу первичной медицинской помощи для молодежи с психическими расстройствами, и пропорционально уменьшилось количество пациентов, привлеченных через специализированную клинику PACE в Orygen. Поскольку клиники headspace имеют более широкий охват, доля пациентов с риском развития психоза в них ниже, чем в специализированной службе PACE, что могло привести к общему снижению частоты развития психоза[77].

Однако временные изменения могут быть связаны не только с исходным риском у пациентов UHR, но и со взаимосвязью между предикторами и клиническими исходами. Например, если изменяются клиническая картина пациентов, направленных в специализированную службу, или стандарты лечения, которое они получают[18,78], это может повлиять на связь между определенными факторами риска и развитием психоза. В таких случаях потребуется обновление или адаптация существующей прогностической модели. Примерами могут служить сдвиг когорты в сторону пациентов с преобладанием характеристик Trait и минимальной выраженностью продуктивных симптомов или внедрение нового стандарта лечения, снижающего влияние продуктивных симптомов на риск развития психоза. Независимый вклад изменений исходного риска, характеристик популяции и тактики ведения пациентов в когорте UHR 1000+ остается предметом для дальнейшего изучения.

Ограничением набора данных, использованного в этом исследовании, является количество пропущенных данных. Оценочные батареи различались во всех включенных исследованиях, что связано с эволюцией знаний с течением времени и ориентацией на конкретные исследовательские вопросы. Несколько методов оценки были заменены или обновлены между первым и последним исследованиями. Стандартизация выборок представляла собой уникальную задачу, охватывающую более двух десятилетий. Более того, анализ выживаемости в этой когорте ограничен различиями периода наблюдения между исследованиями, что может вносить неопределенность и предвзятость в расчетах, использующих информацию о времени до наступления события. Наконец, разделение набора данных по времени сократило размер выборки для разработки и валидации, что потенциально привело к более высокому риску переобучения и менее надежным результатам оценки[79].

В целом набор данных UHR 1000+ представляет собой крупнейшую на сегодняшний день выборку молодых людей с риском развития психоза с периодом наблюдения до 16,7

года и предоставляет уникальный ресурс для дальнейшего изучения моделей прогнозирования исходов группы с ультравысоким риском. Ценность этого ресурса будет далее расширяться за счет интеграции мультимодальных данных, таких как данные нейровизуализации, когнитивной науки, геномики и цифровой оценки.

Наши результаты показывают, что более высокие уровни дезорганизованной речи и необычного содержания мыслей, измеренные с помощью CAARMS, более высокая выраженность негативных симптомов, более низкий уровень социального функционирования и большая продолжительность симптомов до обращения в службу UHR наиболее точно определяют риск развития психоза. Модель, разработанная на когорте UHR 1000+, продемонстрировала умеренную точность (0,65). Однако наши результаты показывают, что крайне важно оценивать существующие и будущие калькуляторы риска на предмет временного отклонения, которое со временем может снизить ценность моделей в клинической практике.

БЛАГОДАРНОСТИ

Эта работа финансировалась через Центр передового опыта в области прогнозирования ранних психических расстройств и профилактического лечения (Prediction of Early Mental Disorder and Preventive Treatment, PRE-EMPT) Национального совета по здравоохранению и медицинским исследованиям (Prediction of Early Mental Disorder and Preventive Treatment – National Health and Medical Research Council Centre of Research Excellence, NHMRC) (грант № 1198304). A. Lin получил поддержку в виде стипендии NHMRC Emerging Leadership Fellowship (№ 2010063); A.R. Yung – стипендий NHMRC Senior Research Fellowship (№ 566593) и Principal Research Fellowship (№ 1136829); P. Amminger – стипендий NHMRC Senior Research Fellowship (№ 1080963); J.T.W. Wigman – гранта Veni от Нидерландской организации научных исследований (NWO) (№ 016.156.019). Дополнительная информация по данному исследованию доступна на https://osf.io/ytjsn.

Перевод: Барканова Валерия, г. Москва
Редактура: Шишковская Татьяна Игоревна, г. Москва

Литература

1. Yung AR, Phillips LJ, McGorry PD et al. Prediction of psychosis: a step towards indicated prevention of schizophrenia. Br J Psychiatry 1998;172(Suppl. 33):14-20.
2. Yung AR, Phillips LJ, Yuen HP et al. Risk factors for psychosis in an ultra highrisk group: psychopathology and clinical features. Schizophr Res 2004;67:13-42.
3. Yung AR, Phillips LJ, Yuen HP et al. Psychosis prediction: 12-month follow up of a high-risk ("prodromal") group. Schizophr Res 2003;60:21-32.
4. Yung AR, McGorry PD. The prodromal phase of first-episode psychosis: past and current conceptualizations. Schizophr Bull 1996;22:353-70.
5. McGorry PD, Yung AR, Phillips LJ. The "close-in" or ultra high-risk model: a safe and effective strategy for research and clinical intervention in prepsychotic mental disorder. Schizophr Bull 2003;29:771-90.
6. Yung AR, McGorry PD. The initial prodrome in psychosis: descriptive and qualitative aspects. Aust N Z J Psychiatry 1996;30:587-99.
7. McGorry PD, Singh BS. Schizophrenia: risk and possibility of prevention. In: Raphael B, Burrows GD (eds). Handbook of preventive psychiatry. Amsterdam: Elsevier, 1995:492-514.
8. Fusar-Poli P, Borgwardt S, Bechdolf A et al. The psychosis high-risk state: a comprehensive state-of-the-art review. JAMA Psychiatry 2013;70:107-20.
9. Salazar de Pablo G, Radua J, Pereira J et al. Probability of transition to psychosis in individuals at clinical high risk: an updated meta-analysis. JAMA Psychiatry 2021;78:970-8.
10. Lee TY, Lee J, Kim M et al. Can we predict psychosis outside the clinical highrisk state? A systematic review of non-psychotic risk syndromes for mental disorders. Schizophr Bull 2018;44:276-85.
11. Coutts F, Koutsouleris N, McGuire P. Psychotic disorders as a framework for precision psychiatry. Nat Rev Neurol 2023;19:221-34.
12. Fusar-Poli P, Salazar de Pablo G, Correll CU et al. Prevention of psychosis: advances in detection, prognosis, and intervention. JAMA Psychiatry 2020;77:755- 65.
13. Sanfelici R, Dwyer DB, Antonucci LA et al. Individualized diagnostic and prognostic models for patients with psychosis risk syndromes: a meta-analytic view on the state of the art. Biol Psychiatry 2020;88:349-60.
14. Dwyer DB, Falkai P, Koutsouleris N. Machine learning approaches for clinical psychology and psychiatry. Annu Rev Clin Psychol 2018;14:91-118.
15. Worthington MA, Cao H, Cannon TD. Discovery and validation of prediction algorithms for psychosis in youths at clinical high risk. Biol Psychiatry Cogn Neurosci Neuroimaging 2020;5:738-47.
16. Studerus E, Ramyead A, Riecher-Rössler A. Prediction of transition to psychosis in patients with a clinical high risk for psychosis: a systematic review of methodology and reporting. Psychol Med 2017;47:1163-78.
17. Moons KGM, Kengne AP, Grobbee DE et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98:691-8.
18. Hartmann JA, Yuen HP, McGorry PD et al. Declining transition rates to psychotic disorder in "ultra-high risk" clients: investigation of a dilution effect. Schizophr Res 2016;170:130-6.
19. Nelson B, Yuen HP, Lin A et al. Further examination of the reducing transition rate in ultra high risk for psychosis samples: the possible role of earlier intervention. Schizophr Res 2016;174:43-9.
20. Yung AR, Yuen HP, Berger G et al. Declining transition rate in ultra high risk (prodromal) services: dilution or reduction of risk? Schizophr Bull 2007;33: 673-81.
21. Spiteri-Staines AE, Yung AR, Lin A et al. Non-psychotic outcomes in young people at ultra-high risk of developing a psychotic disorder: a long-term follow-up study. Schizophr Bull 2024; doi:10.1093/schbul/sbae005.
22. Reynard C, Jenkins D, Martin GP et al. Is your clinical prediction model past its sell by date? Emerg Med J 2022;39:956-8.
23. Hickey GL, Grant SW, Murphy GJ et al. Dynamic trends in cardiac surgery: why the logistic EuroSCORE is no longer suitable for contemporary cardiac surgery and implications for future risk models. Eur J Cardio-Thoracic Surg 2013;43:1146-52.
24. Koutsouleris N, Dwyer DB, Degenhardt F et al. Multimodal machine learning workflows for prediction of psychosis in patients with clinical high-risk syndromes and recent-onset depression. JAMA Psychiatry 2021;78:195-209.
25. Cannon TD, Yu C, Addington J et al. An individualized risk calculator for research in prodromal psychosis. Am J Psychiatry 2016;173: 980-8.
26. Rosen M, Betz LT, Schultze-Lutter F et al. Towards clinical application of prediction models for transition to psychosis: a systematic review and external validation study in the PRONIA sample. Neurosci Biobehav Rev 2021;125:478-92.
27. Studerus E, Beck K, Fusar-Poli P et al. Development and validation of a dynamic risk prediction model to forecast psychosis onset in patients at clinical high risk. Schizophr Bull 2020;46:252-60.
28. Malda A, Boonstra N, Barf H et al. Individualized prediction of transition to psychosis in 1,676 individuals at clinical high risk: development and validation of a multivariable prediction model based on individual patient data meta-analysis. Front Psychiatry 2019;10:345.
29. Fusar-Poli P, Rutigliano G, Stahl D et al. Development and validation of a clinically based risk calculator for the transdiagnostic prediction of psychosis. JAMA Psychiatry 2017;74:493-500.
30. Sox HC, Higgins MC, Owens DK. Measuring the outcome of care – expected utility analysis. In: Sox HC, Higgins MC, Owens DK (eds). Medical decision making, 2nd ed. Chichester: Wiley, 2013:204-42.
31. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Mak 2006;26:565-74.
32. Yung AR, McGorry PD, McFarlane CA et al. The Pace Clinic: development of a clinical service for young people at high risk of psychosis. Australas Psychiatry 1995;3:345-9.
33. Amminger GP, Nelson B, Markulev C et al. The NEURAPRO biomarker analysis: long-chain omega-3 fatty acids improve 6-month and 12-month outcomes in youths at ultra-high risk for psychosis. Biol Psychiatry 2020;87:243-52.
34. McGorry PD, Yung AR, Phillips LJ et al. Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. Arch Gen Psychiatry 2002;59:921-8.

35. Yung AR, Phillips LJ, Nelson B et al. Randomized controlled trial of interventions for young people at ultra high risk for psychosis: 6-month analysis. J Clin Psychiatry 2011;72:430-40.
36. Berger GE, Wood SJ, Ross M et al. Neuroprotective effects of low-dose lithium in individuals at ultra-high risk for psychosis. A longitudinal MRI/MRS study. Curr Pharm Des 2012;18:570-5.
37. Nelson B, Amminger GP, Yuen HP et al. Staged treatment in early psychosis: a sequential multiple assignment randomised trial of interventions for ultra high risk of psychosis patients. Early Interv Psychiatry 2018;12:292-306.
38. McGorry PD, Nelson B, Markulev C et al. Effect of ω-3 polyunsaturated fatty acids in young people at ultrahigh risk for psychotic disorders: the NEURAPRO randomized clinical trial. JAMA Psychiatry 2017;74:19-27.
39. Thompson KN, Phillips LJ, Komesaroff P et al. Stress and HPA-axis functioning in young people at ultra high risk for psychosis. J Psychiatr Res 2007;41:561-9.
40. Phillips LJ, Nelson B, Yuen HP et al. Randomized controlled trial of interventions for young people at ultra-high risk of psychosis: study design and baseline characteristics. Aust N Z J Psychiatry 2009;43:818-29.
41. Bayer JMM, Spark J, Krcmar M et al. The SPEAK study rationale and design: a linguistic corpus-based approach to understanding thought disorder. Schizophr Res 2023;259:80-7.
42. Hartmann JA, Nelson B, Spooner R et al. Broad clinical high-risk mental state (CHARMS): methodology of a cohort study validating criteria for pluripotent risk. Early Interv Psychiatry 2019;13:379-86.
43. European Network of National Networks studying Gene-Environment Interactions in Schizophrenia (EU-GEI). Identifying gene-environment interactions in schizophrenia: contemporary challenges for integrated, large-scale investigations. Schizophr Bull 2014;40:729-36.
44. Tognin S, van Hell HH, Merritt K et al. Towards precision medicine in psychosis: benefits and challenges of multimodal multicenter studies – PSYSCAN: translating neuroimaging findings from research into clinical practice. Schizophr Bull 2020;46:432-41.
45. Nelson B, Lavoie S, Gawęda Ł et al. The neurophenomenology of early psychosis: an integrative empirical study. Conscious Cogn 2020;77:102845.
46. Nelson B, Thompson A, Yung AR. Basic self-disturbance predicts psychosis onset in the ultra high risk for psychosis "prodromal" population. Schizophr Bull 2012;38:1277-87.
47. Overall JE, Gorham DR. The brief psychiatric rating scale. Psychol Rep 1962;10:799-812.
48. Andreasen NC, Flaum M, Arndt S. The Comprehensive Assessment of Symptoms and History (CASH): an instrument for assessing diagnosis and psychopathology. Arch Gen Psychiatry 1992;49:615-23.
49. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed. Washington: American Psychiatric Publishing, 1994.
50. Yung AR, Yung AR, Pan Yuen H et al. Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. Aust N Z J Psychiatry 2005;39:964-71.
51. Goldman H, Skodol A, Lave T. Revising axis V for DSM-IV: a review of measures of social functioning. Am J Psychiatry 1992;149:1148-56.
52. Nelson B, Yuen HP, Wood SJ et al. Long-term follow-up of a group at ultra high risk ("prodromal") for psychosis: the PACE 400 study. JAMA Psychiatry 2013;70:793-802.
53. Andreasen NC. Scale for the Assessment of Negative Symptoms (SANS). Iowa City: University of Iowa, 1983.
54. Samara MT, Engel RR, Millier A et al. Equipercentile linking of scales measuring functioning and symptoms: examining the GAF, SOFAS, CGI-S, and PANSS. Eur Neuropsychopharmacol 2014;24:1767-72.
55. Sterne JAC, White IR, Carlin JB et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.
56. Marshall A, Altman DG, Holder RL et al. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med Res Methodol 2009;9:57.
57. Rubin DB. Multiple imputation. In: van Buuren S (ed). Flexible imputation of missing data, 2nd ed. Boca Raton: Chapman and Hall/CRC, 2018:29-62.
58. Steyerberg EW, Harrell FE, Borsboom GJJM et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54:774-81.
59. Steyerberg EW, Vickers AJ, Cook NR et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128-38.
60. Alba AC, Agoritsas T, Walsh M et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. JAMA 2017;318:1377-84.
61. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. Diagn Progn Res 2019;3:18.
62. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45:1-67.
63. Therneau TM, Grambsch PM. The Cox model. In: Therneau TM, Grambsch PM (eds). Modeling survival data: extending the Cox model. New York: Springer, 2000:39-77.
64. Therneau TM. A Package for survival analysis in R. https://cran.r-project.org/ package=survival.
65. Harrell FE Jr. rms: regression modeling strategies. https://cran.r-project.org/ package=rms.
66. Fusar-Poli P, Cappucciati M, Borgwardt S et al. Heterogeneity of psychosis risk within individuals at clinical high risk: a meta-analytical stratification. JAMA Psychiatry 2016;73:113-20.
67. Mei C, van der Gaag M, Nelson B et al. Preventive interventions for individuals at ultra high risk for psychosis: an updated and extended meta-analysis. Clin Psychol Rev 2021;86:102005.
68. Hartmann S, Cearns M, Pantelis C et al. Combining clinical with cognitive or magnetic resonance imaging data for predicting transition to psychosis in ultra high-risk patients: data from the PACE 400 cohort. Biol Psychiatry Cogn Neurosci Neuroimaging 2024;9:417-28.
69. Dwyer D, Clark S, Amminger GP et al. Towards precision psychiatry for youth: the PRE-EMPT Centre of Research Excellence roadmap. https://www.autho rea.com.
70. Abi-Dargham A, Moeller SJ, Ali F et al. Candidate biomarkers in psychiatric disorders: state of the field. World Psychiatry 2023;22:236-62.
71. Brady LS, Larrauri CA, AMP SCZ Steering Committee. Accelerating Medicines Partnership® Schizophrenia (AMP® SCZ): developing tools to enable early intervention in the psychosis high risk state. World Psychiatry 2023;22:42-3.
72. Chekroud AM, Hawrilenko M, Loho H et al. Illusory generalizability of clinical prediction models. Science 2024;383:164-7.
73. Tanner KT, Keogh RH, Coupland CAC et al. Dynamic updating of clinical survival prediction models in a changing environment. Diagn Progn Res 2023;7:24.
74. Davis SE, Lasko TA, Chen G et al. Calibration drift in regression and machine learning models for acute kidney injury. J Am Med Informatics Assoc 2017;24: 1052-61.
75. Jenkins DA, Martin GP, Sperrin M et al. Comparing predictive performance of time invariant and time variant clinical prediction models in cardiac surgery. Stud Health Technol Inform 2024;310:1026-30.
76. Toll DB, Janssen KJM, Vergouwe Y et al. Validation, updating and impact of clinical prediction rules: a review. J Clin Epidemiol 2008;61:1085-94.
77. McGorry PD, Mei C, Chanen A et al. Designing and scaling up integrated youth mental health care. World Psychiatry 2022;21:61-76.
78. Formica MJC, Phillips LJ, Hartmann JA et al. Has improved treatment contributed to the declining rate of transition to psychosis in ultra-high-risk cohorts? Schizophr Res 2022;243:276-84.
79. Collins GS, Dhiman P, Ma J et al. Evaluation of clinical prediction models (part 1): from development to external validation. BMJ 2024;384:e074819.

DOI:10.1002/wps.21240