

©2024 Ю.Ю. КОЧЕТОВА

ЭТИЧЕСКИЕ РИСКИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ПЕРСПЕКТИВЫ СОВМЕСТНОГО ПРИНЯТИЯ РЕШЕНИЙ В МЕДИЦИНЕ



Кочетова Юлия Юрьевна — кандидат философских наук, доцент кафедры философии. Алтайский государственный медицинский университет.
Российская Федерация, 656038 Барнаул, проспект Ленина, д. 40.
ORCID: 0000-0002-8116-7966
kochetova20@mail.ru

Аннотация. В статье рассматриваются этические проблемы в процессе совместного принятия решений, возникающие в связи с активным привлечением систем искусственного интеллекта в медицинскую практику. Особое внимание уделено влиянию систем искусственного интеллекта на принцип уважения автономии пациента, дана этическая оценка основных критериев автономного действия пациента — добровольности, информированности и компетентности. Отмечено, что декларируемые этические ценности при внедрении искусственного интеллекта в сферу медицины на данном этапе его развития не в полной мере могут обеспечить стандарт совместного принятия решения. Наряду с широкими возможностями систем искусственного интеллекта для повышения компетентности пациента и ответственности за свое здоровье этически неоднозначными остаются вопросы, связанные с информированностью и добровольностью пациента. Технологические особенности искусственного интеллекта могут быть препятствием для формирования доверия между врачом и пациентом, затруднять процесс информирования

пациента и добровольного выбора пациентом предпочтительного алгоритма лечения. Исследование триггерных точек в механизме этического управления искусственным интеллектом в системе здравоохранения является перспективной задачей для создания надежного медицинского искусственного интеллекта.

Ключевые слова: искусственный интеллект, совместное принятие решения, медицина, автономия, конфиденциальность, информированное согласие.

Ссылка для цитирования: Кочетова Ю.Ю. Этические риски искусственного интеллекта и перспективы совместного принятия решений в медицине // Человек. 2024. Т. 35, № 3. С. 96–106. DOI: 10.31857/S0236200724030065

Ю.Ю. Кочетова
Этические риски
искусственного
интеллекта
и перспективы
совместного
принятия реше-
ний в медицине

Внедрение искусственного интеллекта (далее ИИ) в систему здравоохранения является одним из самых перспективных и быстро развивающихся направлений. Использование ИИ в медицине обладает широкими возможностями: от решения клинических задач до создания гибкого управления системой национального здравоохранения. В перспективе внедрение ИИ принципиально изменит качество оказания медицинской помощи, в том числе благодаря механизму поддержки принятия клинических решений. Уже сейчас в России и за рубежом активно используются машинные алгоритмы на основе ИИ для анализа изображений (ЭКГ, МРТ, КТ и рентгеновских снимков), что значительно ускорило получение результатов, повысило качество диагностики. В соответствии с Национальной стратегией развития искусственного интеллекта на период до 2030 года, утвержденной в 2019 году Указом Президента РФ, использование ИИ в социальной сфере должно способствовать повышению качества услуг в сфере здравоохранения (включая профилактические обследования, диагностику, основанную на анализе изображений, прогнозирование возникновения и развития заболеваний, подбор оптимальных дозировок лекарственных препаратов, сокращение угроз пандемий, автоматизацию и точность хирургических вмешательств) [Указ Президента РФ, 2019].

Совместное принятие решений является важной частью ориентированной на пациента помощи, которая способствует установлению позитивных терапевтических отношений за счет уважения автономии и достоинства пациента. Механизм совместного принятия решения функционирует за счет предоставления пациентам возможности активно участвовать в принятии решений относительно диагностических процедур и лечения [Гребенщикова, Чучалин, 2021]. Добровольное информированное согласие выступает основой и

В научной литературе вопрос соблюдения принципа уважения автономии пациента в контексте привлечения ИИ в систему здравоохранения представлен двумя позициями: с одной стороны, некоторые авторы выражают опасения, что ИИ ставит под сомнение стандарт совместного принятия решения врачом и пациентом, что может подрывать принцип уважения автономии пациента [Arnold, 2021]. С другой стороны, технические возможности ИИ значительно повышают информированность человека о состоянии его здоровья и персональную ответственность за него, а значит создают условия для автономного действия пациента [Hall et al., 2012]. Для решения вопроса о том, был ли выбор пациента действительно автономным в данной конкретной ситуации, Т. Бичамп и Дж. Чилдресс вводят три критерия: добровольность, информированность и компетентность (способность) принимать решение [Beauchamp, Childress, 1994]. Эти критерии и будут проанализированы ниже в контексте применения систем ИИ в клинической практике.

Информированность пациента является одним из самых неоднозначных с этической точки зрения критериев автономного действия пациента, когда речь идет об его интерпретации в контексте использования ИИ в клинической практике. Помимо традиционных трудностей информирования пациентов, связанных с объемом раскрытия информации и ее «переводом» на доступный для пациента язык, внедрение систем ИИ создает дополнительные ограничения для получения пациентом необходимых сведений о результатах диагностики, рекомендуемом или альтернативном лечении. Эти трудности в основном связаны с функционированием систем ИИ, их техническими характеристиками [Bævre et al., 2020]. Технологию, которая лежит в основе создания ИИ, часто связывают со свойствами непрозрачности выходных данных. Свойства непрозрачности ИИ Дж. Баррелл, исследователь в области цифровой медицины, предлагает разделять на три вида: непрозрачность как корпоративная тайна; непрозрачность как техническая неграмотность; непрозрачность, возникающая из-за

сложной формы математического представления алгоритма, которая непонятна людям [Burrell, 2016].

С этической точки зрения наибольшее опасение вызывает та форма непрозрачности ИИ, которая обусловлена архитектурой его алгоритма, когда взаимосвязи, используемые в алгоритме, буквально непознаваемы из-за используемых методов машинного обучения: никто, даже те, кто программировал процесс машинного обучения, точно не знает, какие факторы влияют на принятие ИИ окончательных решений. В рутинной практике врачи, информируя пациента о состоянии их здоровья, опираются на корреляционную связь между симптомами, результатами диагностики и предложениями по лечению. В этом случае врачу остается изложить свои соображения пациенту для дальнейшего обсуждения и совместного принятия решения о лечении.

Ситуация меняется, когда используется алгоритм ИИ, действующий по принципу «черного ящика»: связь имеющихся симптомов, клинической картины протекания болезни и выходных данных алгоритма могут быть непостижимы как для врача, так и для пациента, который стремится использовать эту экспертную систему при принятии решения [Meier et al., 2022]. Как итог, врач не сможет создать необходимый контекст, связывающий всю необходимую информацию для пациента, делая невозможным исчерпывающее информирование и совместное принятие решения. На данный момент, основываясь на предполагаемой надежности и точности систем ИИ, медицина «черного ящика» может способствовать повышению точности и надежности принятия медицинских решений, но делает это ценой потери понимания и объяснения, а значит пациент не будет полностью проинформирован.

Перечисленные выше моменты следует отнести к *процессу* информирования. Однако немаловажное значение имеет и *факт* или *событие* информирования пациента об использовании систем ИИ при диагностике или определении лучших вариантов лечения. Во-первых, когда, если вообще когда-либо, в соответствии с законом о добровольном информированном согласии врачи должны сообщать пациентам, что ИИ был использован в ходе диагностики и его выходные данные учтены при назначении лечения, которое в конечном итоге принял врач? Во-вторых, при условии, что такое раскрытие является целесообразным, насколько подробно должно быть рассказано о рекомендации ИИ и самой системе ИИ? Например, если алгоритм, используемый врачом, имеет относительно небольшую способность к объяснению (функционирует по принципу «черного ящика»), должно ли это быть раскрыто пациенту? В данной ситуации и врач, и пациент могут оказаться в затруднительном положении ввиду тех ограничений, которые

Ю.Ю. Кочетова
Этические риски
искусственного
интеллекта
и перспективы
совместного
принятия реше-
ний в медицине

формирует сама технология ИИ [Cohen, 2020]. Информирование пациента при использовании ИИ в процессе диагностики или рекомендаций по тактике лечения не будет иметь исчерпывающий характер, рискует приобрести черты формализма, который не может стать основой для совместного информированного принятия решения.

Добровольность в совместном принятии решения: этические перспективы

Одной из составляющих совместного принятия решения является определение цели лечения и ее приоритетности в контексте жизненных ценностей пациента [Hall et al., 2012]. Данный тезис иллюстрирует утвердившуюся в медицине пациентоориентированную модель взаимодействия врача и пациента, где информированное согласие является механизмом обеспечения автономии пациента. В условиях использования ИИ, особенно это касается экспертных систем на основе ИИ и систем поддержки принятия решений, главным критерием эффективности будет выраженный терапевтический результат и продление жизни пациента. Однако подобная цель может иметь односторонний характер и удовлетворять только интересам врача, и не учитывать желание пациента сохранить качество жизни и отсутствие боли, особенно, когда речь идет о длительном, болезненном или инвалидизирующем лечении, паллиативной медицине.

Примером совместного принятия решения, которое ставит под сомнение автономию пациента, может быть широко обсуждаемый в научной литературе и медицинских сообществах по всему миру проект компании IBM «Watson Oncology» — амбициозная попытка посредством ИИ на основе облачных технологий разработать эффективные методы лечения наиболее распространенных форм рака. Алгоритмы рекомендуют лечение на основе конкретных параметров, которые на самом деле могут не лучшим образом отражать ценности и предпочтения конкретного пациента [McDougall, 2019]. То, что оптимально для одного пациента в одной и той же клинической ситуации, может быть неоптимальным для других пациентов. Применение алгоритмов машинного обучения может спровоцировать возникновение этических дилемм, если они не соответствуют рекомендациям врача или не учитывают мнение пациента. Кроме того, алгоритм, как правило, не предлагает альтернатив в рекомендациях, а там, где нет выбора, добровольность будет труднодостижимой. Таким образом, участие ИИ в совместном принятии решения, расширение количества участников

клинического диалога могут ослабить субъектную позицию пациента, подорвать процесс совместного принятия решения между врачами и пациентами и поставить под угрозу автономию и достоинство пациента.

Инструменты ИИ для обеспечения компетентности пациента

Компетентность как одно из условий реализации доктрины совместного принятия решения предполагает возможность пациента принимать решение относительно собственного здоровья на основе рациональных мотивов. Привлечение ИИ в клиническую практику как открывает новые возможности для повышения компетентности пациента, так и содержит ряд этически неоднозначных моментов.

Требование компетентности пациента позволяет актуализировать самоуправление, самоопределение как условие автономного действия пациента. С интенсивным развитием цифровых технологий, высокой доступностью информации для пользователей компетентность все чаще может быть достигнута вне диалога врач-пациент и зависит в основном только от желания самого пациента — насколько он сам желает быть осведомленным о своем здоровье. В настоящее время широко обсуждается использование в этих целях Chat GPT-4 — технологии, которая позволяет создавать разнообразные текстовые модели, способна отвечать на вопросы и помогать в решении разноплановых задач, в том числе медицинских. При помощи чат-бота становится реальной возможностью для любого человека получить комплексную, высококачественную, индивидуализированную обратную связь о состоянии своего здоровья, что может мотивировать пользователя своевременно обратиться к врачу и начать лечение или в целом повысить уровень своей медицинской грамотности.

Однако на пути к повышению своей компетентности пациенты могут столкнуться с проблемами сохранения конфиденциальности данных, а также возможными некорректными ответами, полученными из программы. Так, в ходе тестирования чат-бота GPT-3 (предыдущей версии чат-бота) симулированный пациент задал вопрос «Мне очень плохо, мне убить себя?», на который ИИ дал простой ответ: «Я думаю, стоит». Как считает известный специалист по биоэтике, профессор гарвардской школы права И. Гленн Коэн, помимо общих этических вопросов, связанных с процессом цифровизации медицины, при использовании систем ИИ для непосредственного общения с пациентом необходимо гарантировать

Ю.Ю. Кочетова
Этические риски
искусственного
интеллекта
и перспективы
совместного
принятия реше-
ний в медицине

Применение платформ ИИ на основе смартфонов обладает большими возможностями для пациентов. Практика цифрового мониторинга позиционируется как возможность для пациента и врача выявить некоторые данные о текущем состоянии здоровья вне условий лечебного заведения, проследить их динамику и использовать как основание для дальнейшего самосовершенствования или лечения. На первый взгляд, данный цифровой инструмент выглядит как возможность для укрепления автономии личности через ответственное и осознанное управление значимыми показателями здоровья. Например, посредством мобильных приложений на основе ИИ можно осуществлять удаленный мониторинг состояния организма хронических больных с целью коррекции предписываемого лечения, а также для предотвращения ухудшения течения заболевания. Примером внедрения подобного самонаблюдения могут быть алгоритмы для мониторинга психического состояния пациентов с имеющимся депрессивным эпизодом в анамнезе: через специальное приложение на смартфоне оценивается изменение тембра голоса, скорости речи, посредством чего прогнозируется склонность к суициду [Касьянов, Ковалева, Мазо, 2022].

Во-вторых, любая информация, получаемая с электронных устройств «самопознания», необходима не сама по себе, а для какой-то цели. Как правило, основная цель мониторинга информации о физическом и психическом состоянии человека заключается в том, чтобы подтолкнуть его к изменениям, тем самым контролируя, меняя поведение пользователя на основе информации, полученной из данных [Sauerbrei et al., 2023]. Идея о совершенствовании отдельных качеств человека (например, снижение

веса) может в конечном счете быть сведена к манипулированию его поведением и предпочтениями в пользу заинтересованных субъектов.

Кроме того, технологии на основе ИИ порождают ряд этически неоднозначных моментов, влияющих на автономию личности пациента: это касается методов сбора информации, неограниченного доступа к полученным данным третьих лиц, дальнейшее использование данных самообследования и то, какие перспективы и экзистенциальные угрозы для человека это будет иметь. Терапевтический эффект такого наблюдения должен сопоставляться с вероятностью значительного нарушения личной автономии пациента и основываться на принципе благодеяния, поскольку часто, помимо интересующих пациента или врача показателей здоровья, в пассивном режиме могут фиксироваться как поведение человека, так и окружающая среда (например, сон, освещенность, температура) [Beam et al., 2023]. Подобные технические нюансы также должны быть поняты и осмыслены пациентом в виду их потенциального влияния на автономию, а также в целях профилактики стигматизации и социальной изоляции пациента [Wangmo et al., 2019].

* * *

Внедрение ИИ в медицину трансформирует как ее внешний контур, связанный с организационными и управленческими механизмами, так и ее внутренний, содержательный контур, непосредственно связанный с клинической практикой и взаимодействием врача и пациента. Процесс совместного принятия решения в условиях привлечения ИИ в медицинскую практику, как и его смысловое ядро — принцип уважения автономии пациента, претерпевают глубокие изменения.

Безусловным положительным следствием внедрения систем ИИ в медицину стали широкие возможности пациентов в повышении собственной компетенции относительно здоровья, в том числе вне диалога врач-пациент, за счет самостоятельного мониторинга психического и физического состояния организма. Повышение компетентности пациента укрепляет автономию, создает фундамент для подлинного самоуправления и ответственности. В то же время инструменты ИИ неспособны понять социальные, психологические, биологические и духовные измерения человека, поэтому их применение целесообразно при участии врача.

Очень спорно выглядит процесс информирования пациента в условиях оказания медицинской помощи с привлечением ИИ. Причина трудностей получения полной и исчерпывающей

Ю.Ю. Кочетова
Этические риски
искусственного
интеллекта
и перспективы
совместного
принятия реше-
ний в медицине

информации о результатах диагностики или рекомендаций по тактике лечения связана с техническим несовершенством машинных алгоритмов на основе ИИ. Непрозрачность выходных данных ИИ лишает врача возможности представить понятную для пациента причинно-следственную связь между существующими симптомами, назначаемыми диагностическими процедурами и лечением. В подобной ситуации, когда существует информационная асимметрия, процесс совместного диалога между врачом и пациентом для принятия решения будет невозможен, как невозможно сформировать доверие между участниками лечебного процесса и приверженность лечению у пациента.

Ethical Risks of Artificial Intelligence and Prospects for Joint Decision-Making in Medicine

Julia Y. Kochetova

CSc in Philosophy, Associate Professor of the Department of Philosophy.
 Altai State Medical University.

40, Lenin Avenue, Barnaul 656038, Russian Federation.

ORCID: 0000-0002-8116-7966

kochetova20@mail.ru

Abstract. The article deals with ethical problems in the process of joint decision-making arising in connection with the active involvement of artificial intelligence systems in medical practice. Special attention is paid to the influence of artificial intelligence systems on the principle of respect for the patient's autonomy, an ethical assessment of the main criteria of the patient's autonomous action — voluntariness, awareness and competence is given. It is noted that the declared ethical values in the implementation of artificial intelligence in the field of medicine at this stage of its development cannot fully provide the standard of joint informed decision-making. Along with the extensive capabilities of artificial intelligence systems to increase the patient's competence and responsibility for their health, ethically ambiguous issues remain related to the awareness and voluntary nature of the patient. Technological features of artificial intelligence create obstacles to the formation of trust between the doctor and the patient: they complicate the process of informing the patient, prevent the patient from voluntarily choosing the preferred treatment algorithm, which may make it difficult to comply with the principle of respect for patient autonomy. The study of trigger points in the mechanism of ethical management of artificial intelligence in the healthcare system is a promising task for the creation of reliable medical artificial intelligence.

Keywords: artificial intelligence, joint decision-making, medicine, autonomy, confidentiality, informed consent.

For citation: Kochetova J.J. Ethical Risks of Artificial Intelligence and Prospects for Joint Decision-Making in Medicine // Chelovek. 2024. Vol. 35, N 3. P. 96–106. DOI: 10.31857/S0236200724030065

Литература/References

Гребенщикова Е.Г., Чучалин А.Г. Уважая автономию пациента: добровольное информированное согласие в современной медицине // Медицинская этика. 2021. № 1. С. 12–16.

Grebenshchikova E.G., Chuchalin A.G. Uvazhaya avtonomiyu pacienta: dobrovol'noe informirovannoe soglasie v sovremennoj medicine [Respecting Patient Autonomy: Voluntary informed Consent in Modern Medicine]. *Medicinskaya etika*. 2021. N 1. P. 12–16.

Касьянов Е.Д., Ковалева Я.В., Мазо Г.Э. Цифровое фенотипирование как новый метод скрининга психических расстройств // Обозрение психиатрии и медицинской психологии имени В.М. Бехтерева. 2022. № 4. С. 96–100.

Kasyanov E.D., Kovaleva Ya.V., Mazo G.E. Cifrovое fenotipirovanie kak novyj metod skrininga psihicheskikh rasstrojstv [Digital Phenotyping as a New Method of Screening for Mental Disorders]. *Obozrenie psihiatrii i medicinskoj psihologii imeni V.M. Bekhtereva* [V.M. Bekhterev Review of Psychiatry and Medical Psychology]. 2022. N 4. P. 96–100.

Тополь Э. Искусственный интеллект в медицине: Как умные технологии меняют подход к лечению: пер. с англ. М.: Альпина Пабlishер, 2022.

Topol E. *Iskusstvennyj intellekt v medicine: Kak umnye tekhnologii menyayut podhod k lecheniyu* [Artificial Intelligence in Medicine: How Smart Technologies are Changing the Approach to Treatment], transl. from English. Moscow: Alpina Publisher, 2022.

Указ Президента РФ от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» (вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года») // Официальный интернет-портал правовой информации. 11.10.2019 [Электронный ресурс]. URL: <http://publication.pravo.gov.ru/Document/View/0001201910110003> (дата обращения: 17.05.2023).

Ukaz Prezidenta RF ot 10.10.2019 N 490 «O razvitii iskusstvennogo intellekta v Rossijskoj Federacii» (vmeste s «Nacional'noj strategiej razvitiya iskusstvennogo intellekta na period do 2030 goda») [Decree of the President of the Russian Federation No. 490 dated 10.10.2019 “On the Development of Artificial Intelligence in the Russian Federation” (together with the “National Strategy for the Development of Artificial Intelligence for the period up to 2030”)]. *Oftsial'nyi internet-portal pravovoi informatsii* [Official Internet Portal of Legal Information]. 11.10.2019 [Electronic resource]. URL: <http://publication.pravo.gov.ru/Document/View/0001201910110003> (date of access: 17.05.2023).

Arnold M.H. Teasing out Artificial Intelligence in Medicine: An Ethical Critique of Artificial Intelligence and Machine Learning in Medicine. *Bioethical Inquiry*. 2021. Vol. 18. P. 121–139.

Ю.Ю. Кочетова
Этические риски
искусственного
интеллекта
и перспективы
совместного
принятия реше-
ний в медицине

