***С.В. Никитенко****, аспирант*

*кафедры международного права*

*СПбГУ*

**ИНФОРМИРОВАННОЕ СОГЛАСИЕ НА МЕДИЦИНСКОЕ**

**ВМЕШАТЕЛЬСТВО С ИСПОЛЬЗОВАНИЕМ СИСТЕМ**

**ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Активное распространение искусственного интеллекта (далее – ИИ) в здравоохранении способствует более эффективной реализации конституционного права человека на охрану здоровья (ст. 41 Конституции РФ[[1]](#footnote-1)) через повышение качества и доступности медицинской помощи[[2]](#footnote-2). Использование новых технологий оказывает существенное влияние на правоотношения между врачом и пациентом, трансформируя классические институты и процедуры, одной из которых является получение согласия пациента на медицинское вмешательство.

Необходимость получения такого согласия является общепризнанным обязательным и предварительным условием для осуществления медицинского вмешательства, которое закреплено как на национальном, так и международном уровнях. Данная процедура выражает принцип автономности пациента[[3]](#footnote-3), несущего ответственность за собственное состояние здоровья. В свою очередь, применение ИИ-систем при вмешательстве существенно затрудняет реализацию механизма получения такого согласия в силу непрозрачности механизма принятия решений и необъяснимости результатов, полученных алгоритмом. Медицинский ИИ, равно как и в любой иной области, является технически сложным продуктом, основанным на труднодоступной для понимания неспециалистом архитектуре и принципах работы. Непрозрачность механизма принятия решений порождает проблемы объяснимости результатов, полученных интеллектуальными системами, а также снижает интерпретируемость[[4]](#footnote-4) таких алгоритмов, что в литературе получило наименование проблемы «чёрного ящика»[[5]](#footnote-5).

На практике, проблема выражается через необходимость предоставления пациенту, не обладающему специальными знаниями, в доступной форме информации о процессе принятия интеллектуальной системой решений при условии их низкой интерпретируемости, вызванной технической сложностью и междисциплинарным характером знаний, положенных в основу соответствующего алгоритма. Кроме того, открытым остаётся вопрос об объёме предоставления информации – будет ли достаточно простого упоминания об использовании ИИ-системы в процессе вмешательства или же необходимо излагать принципы его работы. Ситуация может обостриться до крайности – предоставить обоснование того или иного решения, принятого ИИ, подчас не по силам самим разработчикам[[6]](#footnote-6).

В этом состоит основное различие в традиционном и инновационном способах оказания медицинской помощи: врач может объяснить свои действия и принятые решения[[7]](#footnote-7), а также способен нести за них юридическую, моральную и профессиональную ответственность[[8]](#footnote-8), что в совокупности делает «традиционный» подход более приемлемым для пациентов, как минимум, с психологической точки зрения. С другой стороны, в ущерб своей интерпретируемости ИИ-система способна при правильном применении принимать более точные клинические решения (диагноз, лечение или организационно-управленческие предписания) – данный фактор также может быть решающим при выборе пациентом или специалистом способа медицинского вмешательства.

Не менее важно, что проблемы в объяснении результатов деятельности искусственного интеллекта возникают и со стороны самого медицинского персонала. Врачи и рядовые сотрудники также лишены чёткого представления о механизме работы ИИ-алгоритмов, состоящего из ряда математических формул и закодированных операций, понять смысл которых неспециалисту крайне затруднительно либо и вовсе – невозможно. Выполняя функции оператора ИИ-алгоритма, медицинский работник ограничен способностью загружать входные данные и задавать определённые настройки – механизм и методы принятия решений, будучи изначально заложенными в программу разработчиками, остаются вне контроля и понимания врача, что затрудняет использование технологий в медицинской практике[[9]](#footnote-9). Профессиональная медицинская подготовка позволит отличить неправильные клинические решения, принятые системой, однако, в неочевидных ситуациях способность к критической оценке результатов деятельности носителей ИИ практически нивелируется.

Наконец, предварительное изготовление разработчиками подробной «карты решений», к которым может прийти алгоритм, не применимо к ИИ-технологиям, которые способны самостоятельно, бесконтрольно обучаться и выстраивать собственные логические цепочки и выявлять закономерности[[10]](#footnote-10). Представленный способ решения проблемы интерпретируемости наиболее уместен для т.н. экспертных систем, которые являлись технологическими предшественниками искусственного интеллекта и функционировали исключительно в рамках чётко очерченного набора «сценариев» с заранее определёнными причинно-следственными цепочками (метод «древа решений»): примерами таких систем являются программы «DXplain»[[11]](#footnote-11) и «MYCIN»[[12]](#footnote-12).

Вместе с тем, системы, разработанные на основании указанного метода принятия решений, отличаются низкой точностью полученных результатов в сравнении с методами машинного обучения или глубокими нейронными сетями[[13]](#footnote-13). В связи с чем, нормативно закреплённое требование о необходимости разработки ИИ-систем исключительно на архитектуре «древа решений» также не является способом решения проблемы, поскольку это существенно снизит эффективность применения технических решений и может в целом привести к стагнации отрасли по причине функциональной ограниченности метода.

Как результат, проблема «чёрного ящика» усложняется механизм получения информированного согласия пациента на медицинское вмешательство с использованием ИИ-технологий: пациент не сможет принять обоснованное решение без понимания способа осуществления вмешательства и возможных рисках для своего здоровья. В рамках представленной работы, мы предприняли попытку улучшить существующую процедуру получения согласия на медицинское вмешательство, предусмотренную российским законодательством, в целях учёта особенностей применения ИИ-систем, а также обеспечения прав и законных интересов пациента.

Вопрос получения согласия пациента на вмешательство урегулирован в Федеральном законе от 21.11.2011 № 323-ФЗ «Об основах охраны здоровья граждан в Российской Федерации»[[14]](#footnote-14) (далее – Закон № 323-ФЗ), согласно ч. 1 ст. 20 которого информированное добровольное согласие гражданина на медицинское вмешательство является необходимым предварительным условием такого вмешательства и должно быть дано (1) добровольно (2) на основании предоставленной медицинским работником (3) в доступной форме (4) полной информации (5) о целях, методах оказания медицинской помощи, связанном с ними риске, (6) возможных вариантах медицинского вмешательства, о его последствиях, а также (7) о предполагаемых результатах оказания медицинской помощи. В случае отказа от вмешательства лицу также должны в доступной для него форме разъяснить возможные последствия такого отказа (ч. 4 ст. 20 Закона № 323-ФЗ).

Вместе с тем, при анализе Перечня определенных видов медицинских вмешательств, на которые граждане дают информированное добровольное согласие при выборе врача и медицинской организации для получения первичной медико-санитарной помощи, утв. Приказом Минздравсоцразвития России от 23.04.2012 № 390н[[15]](#footnote-15), мы не заметим упоминаний использования интеллектуальных систем в качестве отдельного вида вмешательства или его составной части. Перечисленные в указанном акте виды вмешательств выделены в зависимости от их предметной области и формы, к примеру, «неинвазивные исследования органа слуха» и «исследование функций нервной системы». В то же время, способ вмешательства не выделяется, что делает некорректным с точки зрения юридической техники расширение указанного Перечня за счёт включения в него указаний на ИИ-системы.

Обратимся к международному опыту регулирования в рамках рассматриваемой проблематики. Так, Всеобщая декларация о биоэтике и правах человека, принятая на 18-м пленарном заседании Генеральной конференции ЮНЕСКО 19.10.2005[[16]](#footnote-16), являясь международным договором рекомендательного характера на глобальном международном уровне, также достаточно подробно регулирует механизм и условия дачи согласия на вмешательство. Так, согласно п. 1 ст. 6 Декларации медицинское вмешательство в профилактических, диагностических и терапевтических целях должно осуществляться только с предварительного, свободного, явно выраженного и информированного согласия лица на основе надлежащей информации. Указывается право лица отозвать согласие в любое время и по любой причине без негативных последствий или ущерба.

Примечательно, что в п. 2 данной статьи механизм получения согласия также распространяется на проведение научных исследований, при этом добавляется такое условие, как адекватность и понятная форма информации, на основе которой лицо принимает решение. Полагаем, в данной ситуации указанные условия по смыслу эквивалентны требованию «надлежащего» характера информации из п. 1 ст. 6 Декларации. Данное дополнение имеет рациональное основание при условии чёткого разграничения понятий медицинского вмешательства и научных исследований.

На международном региональном уровне вопрос получения информированного согласия находит отражение в Конвенции Совета Европы о защите прав и достоинства человека в связи с применением достижений биологии и медицины 1997 г.[[17]](#footnote-17), в ст. 5 которой указано общее правило: медицинское вмешательство может осуществляться лишь после того, как лицо, которому предоставлена информация о цели, характере последствиях и рисках вмешательства, даст на это свое добровольное информированное согласие. Россия не подписала и не ратифицировала данную Конвенцию, однако, как мы можем видеть, представленное в ст. 20 Закона № 323-ФЗ положение об информированном согласии является более развёрнутым и полностью включает в себя все условия, перечисленные в Конвенции.

На основе анализа представленных правовых актов государственного и международного уровня мы можем утверждать, что вопрос об особенностях предоставления информированного согласия для вмешательства с использованием ИИ-технологий специальным образом не урегулирован. По данной причине остаются нерешёнными проблемы, вызванные низкой интерпретируемостью интеллектуальных систем и следующей из этого сложностью в предоставлении адекватной, полной информации в доступной форме о предстоящем вмешательстве.

Исходя из уровня технического исполнения современных носителей искусственного интеллекта, с учётом отсутствия полноценного регулирования отрасли и несформировавшегося отношения к технологии со стороны специалистов и пациентов, полагаем обоснованным включить в информированное согласие указание на использование ИИ-алгоритмов при осуществлении определённого вида медицинского вмешательства либо включить общее положение о согласии на использование искусственного интеллекта при оказании медицинской помощи (в виде отдельного поля в бланке согласия на вмешательство или отдельного документа). Закономерным образом, в случае отказа в предоставлении согласия, пациенту должна быть предложена медицинская помощь, оказываемая без использования автоматизированных алгоритмов принятия клинических решений.

Реализация указанных условий механизма информированного согласия непосредственно связана с принятием комплекса мер по преодолению проблемы «черного ящика», в основе которой лежит полнота и доступность информационных процессов, обрабатываемых ИИ-системами. Приведём ряд мер, которые, по нашему мнению, будут в наибольшей мере способствовать повышению объяснимости и интерпретируемости интеллектуальной модели.

Полагаем, что преодолеть проблему низкой объяснимости (интерпретируемости) результатов работы интеллектуальных систем позволит следующий комплекс мер. Так, ключевое значение для человека, оценивающего результаты «мыслительной» деятельности ИИ-технологий для принятия собственного решения, имеет понимание (1) всех способов решения поставленной клинической задачи, (2) степень эффективности применения каждого из таких способов (вероятность успешного завершения операции или уровень точности диагноза), (3) основные факторы, влияющие на эффективность конкретного способа, а также (4) несколько главных причин возможного «негативного» результата. По нашему мнению, предоставление указанной информации является достаточным для соблюдения требований полноты и доступной формы сведений в целях получения информированного согласия – пациенту и врачу не обязательно знать все аспекты технической стороны работы алгоритма; более того, это практически не достижимо.

Повышению прозрачности будет способствовать предъявление к разработчикам требования о включении в ИИ-программу функции формирования схематичного описания цепочки операций, которые привели к полученному результату. К примеру, такое описание может быть представлено в виде блок-схемы, демонстрирующей все вариации (древо) решений, к которым алгоритм мог прийти на определённом этапе своих вычислений, а также аргументацию, почему был выбран конкретный вариант и отклонены иные. Подобная функция реализована разработчиками системы для обработки генетической информации «Emedgene»[[18]](#footnote-18). При этом, должен быть найден баланс между упрощением в объяснении технической части алгоритма и сохранением как можно более полного изложения принципов работы технологии и факторов, которые могут повлиять на полученные результаты.

В научной литературе предлагается иной способ интерпретации полученных ИИ-системой результатов через предоставление врачу наиболее близкого аналога изображения или описания патологии из имеющейся верифицированной базы данных. Поскольку специалисту для сравнения выводится уже проверенный пример клинического решения, основанный на аналогичных исходных данных, имеется возможность, по крайней мере, сверить оба результата: от ИИ и из базы данных – на предмет наличия соответствующих аномалий и правильности их «квалификации» системой[[19]](#footnote-19).

Вместе с тем, данный способ обеспечения объяснимости результатов применим только для ИИ-алгоритмов, основанных на методе т.н. сиамских нейронных сетей[[20]](#footnote-20). В связи с чем, представленный подход не является универсальным и не может быть закреплён в качестве общего требования для различных интеллектуальных систем: к примеру, алгоритмы, основанные на анализе статистических моделей и больших данных, в принципе не смогут «проиллюстрировать» своё решение набором из одного или нескольких примеров из баз данных, поскольку обнаруживаемые ими закономерности могут быть видны только в масштабе тысяч образцов.

Кроме того, учитывая разные уровни подготовки участников отношений, следует предоставлять две версии соответствующего отчёта: очевидно, для специалиста должен быть сформирован более детализированный отчёт с использованием медицинской терминологии, что позволит более обоснованно подойти к оценке деятельности ИИ-системы. С этим обстоятельством тесно связано косвенное направление для решения проблемы объяснимости искусственного интеллекта: повышение квалификации медицинских кадров в части обучения навыкам работы с большими данными, интеллектуальными системами и интерпретации полученных ими решений.

Указанная функция формирования отчётности должна быть «по умолчанию» встроена во все интеллектуальные алгоритмы и выступить составной частью концепции «protection by design», которую следует внедрить в ИИ-отрасль. Наличие возможности предоставить программой отчёт о совершённых либо предстоящих действиях является одним из ключевых элементов в обеспечении прозрачности и интерпретируемости ИИ-систем, что повысит уровень доверия к технологии и темпы её развития за счёт повышения эффективности механизма «обратной связи».

На наш взгляд, закрепление изложенных рекомендаций в нормативно-правовых актах и стандартах, в том числе, международного уровня позволит не только адаптировать механизм получения информированного согласия пациента на медицинское вмешательство для случаев оказания медицинской помощи с использованием ИИ-алгоритмов, но также решит проблему «чёрного ящика» благодаря обеспечению объяснимости интеллектуальных систем. В то же время, объяснимость как свойство интеллектуальной системы должно стать одним из наиболее важных элементов правового режима технологии, который лежит в основе решения ряда принципиальных для отрасли проблем, связанных, к примеру, с ответственностью за вред, причинённый ИИ-системой при оказании медицинской помощи.

1. Конституция Российской Федерации: принята всенародным голосованием 12.12.1993, с изменениями, одобренными в ходе общероссийского голосования 01.07.2020 [Электронный ресурс] // Справочная правовая система «Консультант-Плюс». Режим доступа: <http://www.consultant.ru/>. [↑](#footnote-ref-1)
2. Привалов С.А. Технологии искусственного интеллекта в сфере обеспечения права на охрану здоровья, доступную и качественную медицинскую помощь: перспективы и проблемы регулирования // Вестник СГЮА. 2021. № 4 (141). С. 36 ; Thrall J. H., Li X., Li Q., Cruz C., Do S., Dreyer K., Brink J. Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success // Journal of the American College of Radiology. Т. 15. № 3. С. 504—508. [↑](#footnote-ref-2)
3. Киселева А.Ю. Применение искусственного интеллекта в здравоохранении: аспекты медицинского права // Медицинское право: теория и практика. 2020. Т. 6. № 2 (12). С. 30. [↑](#footnote-ref-3)
4. Ching T., Himmelstein D.S., Beaulieu-Jones B.K., Kalinin A.A., Do B.T., Way G.P., et al. Opportunities and obstacles for deep learning in biology and medicine [Электронный ресурс] // Journal of the Royal Society Interface. 2018. № 15 (141). 47 с. DOI: 10.1098/rsif.2017.0387 [↑](#footnote-ref-4)
5. Ившин А.А., Гусев А.В., Новицкий Р.Э. Искусственный интеллект: предиктивная аналитика перинатального риска // Вопросы гинекологии, акушерства и перинатологии. 2020. № 19(6). С. 136 ; Carter M.S., Rogers W., Win K.T., Frazer H., Richards B., Houssami N. The ethical, legal and social implications of using artificial intelligence systems in breast cancer care // The Breast. 2020. № 49. С. 26. DOI: 10.1016/j.breast.2019.10.001 ; Antoniades Ch., Asselbergs F.W., Vardas P. Год в сердечно-сосудистой медицине, 2020 г.: цифровое здравоохранение и инновации // Российский кардиологический журнал. 2021. № 26 (3). С. 122. [↑](#footnote-ref-5)
6. Киселева А.Ю. Применение искусственного интеллекта в здравоохранении: аспекты медицинского права // Медицинское право: теория и практика. 2020. Т. 6. № 2 (12). С. 30-31. [↑](#footnote-ref-6)
7. Kamensky S. Artifificial Intelligence and Technology in Health Care: Overview and Possible Legal Implications // DePaul journal of health care law. 2020. Т. 21. № 3. С. 3. [↑](#footnote-ref-7)
8. Иванова А.П. Правовые проблемы использования искусственного интеллекта в сфере здравоохранения // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Серия 4: государство и право. Реферативный журнал. 2021. № 1. С. 153. [↑](#footnote-ref-8)
9. Holzinger A., Biemann C., Pattichis C.S., Kell D.B. What do we need to build explainable AI systems for the medical domain? [Электронный ресурс] // arXiv:1712.09923 [cs.AI]. 2017. [↑](#footnote-ref-9)
10. Киселева А.Ю. Применение искусственного интеллекта в здравоохранении: аспекты медицинского права // Медицинское право: теория и практика. 2020. Т. 6. № 2 (12). С. 31. [↑](#footnote-ref-10)
11. Using decision support to help explain clinical manifestations of disease [Электронный ресурс] // MGHLCS.ORG – официальный сайт Лаборатории компьютерных наук Массачусетской больницы общего профиля. URL: <http://www.mghlcs.org/projects/dxplain> (дата обращения: 01.02.2022). [↑](#footnote-ref-11)
12. Shortliffe E.H., et al. Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system // Computers and biomedical research. 1975. Т. 8. № 4. С. 303-320. [↑](#footnote-ref-12)
13. Мелдо А.А., Уткин Л.В., Трофимова Т.Н., 2020 г. Искусственный интеллект в медицине: современное состояние и основные направления развития интеллектуальной диагностики // Лучевая диагностика и терапия. 2020. Т. 11. № 1. С. 15. [↑](#footnote-ref-13)
14. Об основах охраны здоровья граждан в Российской Федерации: Федеральный закон от 21.11.2011 № 323-ФЗ, ред. от 02.07.2021 [Электронный ресурс] // Справочная правовая система «Консультант-Плюс». Режим доступа: <http://www.consultant.ru/>. [↑](#footnote-ref-14)
15. Об утверждении Перечня определенных видов медицинских вмешательств, на которые граждане дают информированное добровольное согласие при выборе врача и медицинской организации для получения первичной медико-санитарной помощи: Приказ Минздравсоцразвития России от 23.04.2012 № 390н [Электронный ресурс] // Справочная правовая система «Консультант-Плюс». Режим доступа: <http://www.consultant.ru/>. [↑](#footnote-ref-15)
16. Всеобщая декларация о биоэтике и правах человека: принята резолюцией Генеральной конференции ЮНЕСКО по докладу Комиссии III на 18-м пленарном заседании 19.10.2005 [Электронный ресурс] // Официальный сайт ООН. URL: <https://www.un.org/ru/documents/decl_conv/declarations/bioethics_and_hr.shtml> (дата обращения: 13.02.2022). [↑](#footnote-ref-16)
17. Конвенция о защите прав человека и человеческого достоинства в связи с применением достижений биологии и медицины («Конвенция о правах человека и биомедицине»): Совет Европы, ETS № 164, заключена в г. Овьедо 04.04.1997 [Электронный ресурс] // Справочная правовая система «Консультант-Плюс». Режим доступа: <http://www.consultant.ru/>. [↑](#footnote-ref-17)
18. Emedgene. A genomics intelligence platform using artificial intelligence [Электронный ресурс] // URL: <https://www.emedgene.com> (дата обращения: 02.02.2022). [↑](#footnote-ref-18)
19. Мелдо А.А., Уткин Л.В., Трофимова Т.Н., 2020 г. Искусственный интеллект в медицине: современное состояние и основные направления развития интеллектуальной диагностики // Лучевая диагностика и терапия. 2020. Т. 11. № 1. С. 15. [↑](#footnote-ref-19)
20. Bromley J., Bentz J.W., Bottou L., Guyon I., LeCun Y., Moore C., Sackinger E., Shah R. Signature verification using a Siamese time delay neural network // International Journal of Pattern Recognition and Artificial Intelligence. 1993. № 7 (4). С. 737–744 ; Koch G., Zemel R., Salakhutdinov R. Siamese neural networks for one-shot image recognition // Proceedings of the 32nd International Conference on Machine Learning. 2015. № 37. С. 1–8. [↑](#footnote-ref-20)