

АНДРЕЙЧЕНКО А.Е.,

к.ф.-м.н., ООО «К-СКАЙ», г. Петрозаводск, Россия, e-mail: aandreychenko@webiomed.ru

КАФТАНОВ А.Н.,

к.м.н., ООО «К-СКАЙ», г. Петрозаводск, Россия, e-mail: akaftanov@webiomed.ru

ГУСЕВ А.В.,

к.т.н., ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения» Минздрава России, г. Москва, Россия, e-mail: agusev@webiomed.ai

МЕТОДИКА ОЦЕНКИ КАЧЕСТВА ДАННЫХ ЭЛЕКТРОННЫХ МЕДИЦИНСКИХ КАРТ

DOI: 10.25881/18110193_2024_4_48

Аннотация. Ключевой функциональной возможностью медицинских информационных систем медицинских организаций является ведение электронных медицинских карт (ЭМК), которые играют неотъемлемую роль в современной практике здравоохранения, позволяя медицинским организациям последовательно собирать, систематизировать и предоставлять медицинским работникам доступ к информации о диагностике и лечении пациентов. Несмотря на наличие большого объема накопленных ЭМК и давнюю историю их разработки и развития, современные ЭМК имеют довольно низкое качество собираемой в них клинической информации. В настоящий момент нет рекомендуемого подхода к оценке качества данных ЭМК.

Цель. Разработать методику оценки качества данных, содержащихся в ЭМК.

Материалы и методы. Были собраны и систематизированы требования к процедуре оценки качества данных ЭМК и расчета индекса качества по итогам такой процедуры. На основе требований была сформирована методика оценки качества данных, для каждого из этапов методики проработаны подходы ее практической реализации и приведены конкретные примеры расчетов критериев качества для самых распространенных базовых элементов ЭМК.

Результаты. В работе представлена методика оценки качества данных ЭМК, а также алгоритм расчета итоговых индексов качества на основе данных платформы Webiomed. Методика позволяет получить не только интегральную оценку качества, но и ее составляющие, оценивающие разные параметры качества данных, а также детализировать оценку качества по разным элементам ЭМК.

Заключение. Разработанная методика позволяет оценить базовые элементы ЭМК. Также предлагаемая методика предоставляет подход и алгоритм расширения на любые дополнительные элементы ЭМК.

Ключевые слова: методика, электронная медицинская карта, ЭМК, качество данных, оценка качества данных.

Для цитирования: Андрейченко А.Е., Кафтанов А.Н., Гусев А.В. Методика оценки качества данных электронных медицинских карт. Врач и информационные технологии. 2024; 4: 48-59. doi: 10.25881/18110193_2024_4_48.

ANDREYCHENKO A.E.,

PhD, K-Skai, Petrozavodsk, Russia, e-mail: aandreychenko@webiomed.ru

KAFTANOV A.N.,

PhD, K-Skai, Petrozavodsk, Russia, e-mail: akaftanov@webiomed.ru

GUSEV A.V.,

PhD, Federal Research Institute for Health Organization and Informatics, Moscow, Russia,
e-mail: agusev@webiomed.ai

METHODOLOGY FOR ASSESSING THE QUALITY OF ELECTRONIC MEDICAL RECORDS DATA

DOI: 10.25881/18110193_2024_4_48

Abstract. *Background.* The key functionality of the medical information system (MIS) is the maintenance of electronic medical records (EMR), which play an integral role in modern healthcare practice, allowing medical organizations to consistently collect, systematize and provide medical professionals with access to information on the diagnosis and treatment of patients. Despite the existence of a large volume of accumulated EMRs and a long history of their design and development, modern EMRs have a rather low quality of clinical information collected in them. Currently, there is no recommended approach to assessing the quality of EMR data.

Objective. To develop a methodology for assessing the quality of data contained in EMRs.

Materials and methods. Requirements for the procedure of EMR data quality assessment and calculation of the quality index based on the results of such procedure were collected and systematized. Based on the requirements, a methodology for data quality assessment was formed, approaches to its practical implementation were worked out for each stage of the methodology and specific examples of quality criteria calculations for the most common basic elements of EMC were given.

Results. The paper presents a methodology for assessing the quality of EMR data, as well as an algorithm for calculating the final quality indices based on the Webiomed platform data. The methodology allows us to obtain not only an integral quality assessment, but also its components assessing different data quality parameters, as well as to detail the quality assessment for different EHRs elements.

Conclusion. The developed methodology allows to evaluate the basic elements of the EHRs. The proposed methodology also provides an approach and an algorithm for extending to any additional element of the EHRs.

Keywords: methodology, electronic health record, EHR, data quality, data quality assessment.

For citation: Andreychenko A.E., Kaftanov A.N., Gusev A.V. Methodology for assessing the quality of electronic medical records data. Medical doctor and information technology. 2024; 4: 48-59. doi: 10.25881/18110193_2024_4_48.

ВВЕДЕНИЕ

К настоящему времени в России осуществлено полномасштабное внедрение медицинских информационных систем (МИС) медицинских организаций (МО). Свыше 90% государственных МО используют различные МИС [1].

Ключевой функциональной возможностью МИС МО является ведение электронных медицинских карт (ЭМК), которые играют неотъемлемую роль в современной практике здравоохранения, позволяя МО последовательно собирать, систематизировать и предоставлять медицинским работникам доступ к информации о диагностике и лечении пациентов. Ведение медицинскими работниками ЭМК привело к накоплению огромных объемов оцифрованных медицинских данных, которые являются оцифрованным артефактом реальной клинической практики и обладают большим потенциалом для цифровой трансформации отрасли, в том числе внедрения систем поддержки принятия врачебных решений (СППВР), управления на основе данных, исследований реальной клинической практики (RWD) и т.д.

Несмотря на наличие большого объема накопленных ЭМК и давнюю историю их разработки и развития, современные ЭМК имеют довольно низкое качество собираемой в них клинической информации [2]. В литературе есть работы по количественной оценке качества данных ЭМК, однако они не дают полноценного руководства для практического применения, либо требуют расчета очень большого количества метрик качества [3–5]. Так в работе Jiayin Z. et al. (2024) авторы предлагают использование 33 метрик, при этом не дают руководства по их расчету [5]. В этой связи необходимо создание единой, простой в применении и интерпретируемой методики оценки качества данных ЭМК и расчета «Индекса качества ЭМК», что и явилось целью данного исследования.

МАТЕРИАЛЫ И МЕТОДЫ

Разработка методики оценки качества данных ЭМК и расчета соответствующего «Индекса качества ЭМК» выполнялась в несколько этапов. На первом этапе были сформулированы базовые функциональные требования, в том числе путем анализа запросов пользователей платформы Webiomed. На втором этапе на основе

данных литературы было проведено исследование с целью определения основных критериев, по которым необходимо оценивать содержимое ЭМК. Для поиска релевантных источников информации были использованы 4 электронные библиографические базы данных: PubMed, Web of Science, Scopus и РИНЦ. В поисковом запросе применялись строки «Data quality assessment of Electronic health records» и «Оценка качества данных электронных медицинских карт». Период поиска: все публикации до 01.03.2024. Было обнаружено 37 публикаций. После анализа названий и резюме для пристального изучения отобрано 5 публикаций по тематике методики оценки качества данных ЭМК. На третьем этапе была проведена разработка алгоритма расчета критериев и итогового Индекса качества ЭМК, удовлетворяющего базовым функциональным требованиям, сформулированным на первом этапе. Наконец, была разработана практическая реализация методики оценки качества данных для базовых универсальных элементов ЭМК.

РЕЗУЛЬТАТЫ

Сформулированы следующие базовые функциональные требования к Индексу качества ЭМК, рассчитываемому с помощью разрабатываемой методики:

1. Индекс качества ЭМК должен быть количественным.
2. Значения индекса должны находиться в диапазоне 0–100.
3. В первую очередь качество ЭМК должно быть оценено с точки зрения ее пригодности для использования в СППВР, в этой связи необходимо оценить клиническую информацию, содержащуюся в записях ЭМК, в том числе с учетом наличия минимально необходимых для работы СППВР данных.
4. В алгоритме расчета индекса необходимо учесть полноту заполнения социальных данных (пол, дата рождения, прикрепление, участок/МО прикрепления, район проживания).
5. В алгоритме расчета индекса необходимо предусмотреть корректный учет открытых случаев лечения, которые при их открытии могут быть мало заполнены, но затем могут быть полностью заполнены на момент закрытия случая.

6. Необходимо отслеживание динамики индекса качества ЭМК для случаев лечения от их открытия до закрытия.
7. Необходимо предусмотреть проверку на присутствие шаблонных неинформативных текстов в документах ЭМК.

По результатам исследования литературы была выявлена следующая вариативность подходов к выбору критериев при оценке качества данных ЭМК. В 2013 году Weiskopf предложил следующие параметры качества данных ЭМК [6]:

1. Корректность: правдивость данных в ЭМК;
2. Согласованность: соответствие между элементами внутри ЭМК и между ЭМК и другими источниками данных;
3. Правдоподобность: степень, в которой данные ЭМК имеют смысл в широком медицинском контексте;
4. Актуальность: точность данных ЭМК в тот период времени, когда они были записаны, и насколько актуальны сейчас эти данные.

В 2016 году Kahn была проведена

гармонизация параметров и выделены 3 параметра [7]:

1. Согласованность;
2. Полнота;
3. Правдоподобность.

Проанализировав эти и другие научные данные и рекомендации, нами была предложена собственная методика оценки качества ЭМК, включающая учет следующих критериев:

- Полнота: наличие и заполненность обязательных элементов ЭМК;
- Корректность: допустимость формата и значений данных, содержащихся в ЭМК;
- Правдоподобность: соответствие значений элементов и временных рамок их получения в медицинской практике и контексту.

Более подробно результаты нашего исследования подходов к выбору критериев оценки качества данных ЭМК представлены в обзоре [8].

Выбранные критерии качества данных ЭМК для включения в методику оценки представлены в Таблице 1 вместе с методом их расчета.

Таблица 1 — Критерии качества данных ЭМК и методы их оценки

ID	Критерий качества ЭМК	Метод оценки
A	Полнота: наличие и заполненность обязательных элементов ЭМК	
A1	Наличие элемента	Проверка наличия и заполненности (не пустое) в ЭМК следующих элементов:
A2	Заполненность элемента	1. Социально-демографические данные (пол, возраст и т.д.) 2. Минимально-необходимые клинические и антропометрические данные (АД, рост, вес и т.д.) 3. Наличие данных о случаях обращения за медицинской помощью (законченные случаи, истории болезни и т.д.) 4. Наличие минимально-необходимых медицинских записей (врачебный осмотр, протокол инструментального обследования и т.д.)*
B	Корректность: допустимость значений признаков и элементов, содержащихся в ЭМК	
B1	Корректность элементов ЭМК	Проверка корректности полей в случаях, корректность полей в структурированных электронных медицинских документах.**
B2	Корректность значений признаков	Проверка на количество ошибочных признаков, выявленных после формато-логического контроля (ФЛК)
C	Правдоподобность: соответствие значений признаков и значений элементов временным рамкам их получения в медицинской практике и контексту	
C1	Значение (вне времени)	Логический контроль на основе установленных правил (например, постоянство роста у взрослых, гендерные признаки, показатели, заболевания)***
C2	Актуальность (соответствие временных рамок)	Логический контроль на основе установленных правил (например, дата открытия случая раньше даты закрытия случая)***

Примечание: * — перечень элементов ЭМК приведен в Таблице 2. ** — некоторые правила проверки корректности элементов приведены в Таблице 3. *** — перечень правил проверки значений и актуальности приведен в Таблице 4.

Для определения критерия «Полнота: наличие и заполненность обязательных элементов ЭМК» необходимо, чтобы платформа либо МИС принимала, поддерживала и хранила следующие элементы, представленные в таблице 2.

Для определения критерия «Корректность заполнения элементов ЭМК» необходимо разработать следующие опорные словари:

1. Допустимые форматы данных для каждого элемента ЭМК.

2. Пополняемый словарь шаблонных текстов.

В случае заполненности данного элемента ЭМК происходит двухэтапная проверка на соответствие формата данных, хранимых в элементе ЭМК, затем на отсутствие шаблонного текста. Проверка на отсутствие шаблонного текста заканчивается положительным результатом только в случае если текст элемента и шаблонного текста на 100% совпадают.

Итоговый результат проверки корректности заполнения элементов ЭМК является

Таблица 2 — Перечень базовых элементов ЭМК

№	Элемент	Описание
	Социально-демографические данные	В ЭМК должны быть заполнены следующие поля: 1. ФИО или хеш ФИО (закодированное или зашифрованное значение поля) 2. СНИЛС или хеш СНИЛСа 3. Данные о страховом полисе или хеш этих данных 4. Пол 5. Дата рождения 6. Регион проживания (прикрепления) 7. МО прикрепления
	Минимально-необходимые клинические, антропометрические и анамнестические данные	В ЭМК должны быть внесены следующие данные: 1. АД 2. Рост 3. Вес 4. ЧСС 5. Общий холестерин 6. Глюкоза 7. Лист окончательных диагнозов (должна быть хотя бы 1 запись) 8. Пол 9. Частота дыхания 10. Табакокурение 11. Креатинин 12. Гемоглобин 13. Температура
	Данные о случаях обращения за медицинской помощью	В ЭМК пациента должна содержаться хотя бы 1 запись о случае обращения пациента в МО, включая законченный случай, или история болезни, или случай обращения с профилактическими целями (карта диспансеризации, лист вакцинаций и т.д.).
	Минимально-необходимые медицинские записи	В ЭМК пациента должны содержаться хотя бы по 1 экземпляру следующие документы: 1. Протокол инструментального исследования, в котором заполнено «Заключение» 2. Протокол лабораторного исследования, в котором есть хотя бы 1 заполненный параметр (тест) 3. Протокол врачебного осмотра, в котором заполнены поля «Жалобы», «Объективные данные» и указан код окончательного клинического диагноза

Таблица 3 — Некоторые правила оценки корректности заполнения элементов ЭМК

№	Элемент	Поле	Допустимые форматы	Шаблонные тексты
1	Социально-демографические данные	ФИО или хеш ФИО (закодированное или зашифрованное значение поля)	*	N/A
		СНИЛС или хеш СНИЛСа	*	N/A
		Данные о страховом полисе или хеш этих данных	*	N/A
		Пол	*	N/A
		Дата рождения	*	N/A
		Регион проживания (прикрепления)	*	
		МО прикрепления	*	
2	Минимально-необходимые клинические, антропометрические и анамнестические данные	АД	*	N/A
		Рост	*	N/A
		Вес	*	N/A
		ЧСС	*	N/A
		Общий холестерин	*	N/A
		Глюкоза	*	N/A
		Лист окончательных диагнозов	Диагноз соответствует коду МКБ-10 и имеет дату регистрации	N/A
		Частота дыхания		N/A
		Табакокурение		N/A
		Креатинин		N/A
3	Данные о случаях обращения за медицинской помощью	Гемоглобин		N/A
		Температура		N/A
4	Минимально-необходимые медицинские записи	В ЭМК пациента должна содержаться хотя бы 1 запись о случае обращения пациента в МО, включая законченный случай, или история болезни, или случай обращения с профилактическими целями (карта диспансеризации, лист вакцинаций и т.д.). Не учитываются «Разовые посещения» – это не реальный случай обращения, а искусственный компонент цифрового профиля пациента в Webiomed.	Перечень обязательных полей случая и допустимые форматы их наполнения	Проверка на шаблонность и неинформативность случаев лечения
		Протокол инструментального исследования, в котором заполнено «Заключение»	*	Список шаблонных заключений
		Протокол лабораторного исследования, в котором есть хотя бы 1 заполненный параметр (тест)	*	N/A
		Протокол врачебного осмотра, в котором заполнены поля «Жалобы», «Объективные данные» и указан код окончательного клинического диагноза	*	Список шаблонных текстов «Жалоб», «Объективные данные»

Примечание: * — на основе словаря допустимых форматов полей. N/A — неприменимо к данному полю элемента.

положительным (проверка пройдена), если первый этап выявил соответствие формата, а проверка шаблонности дала отрицательный результат.

Для определения критерия «Правдоподобность значений признаков» необходимо сформулировать правила для недопустимых значений признаков.

Правила делятся на две категории:

1. Статичные правила определяют рамки возможных изменений значений полей элементов и признаков без привязки к дате, событию (например, постоянство роста, веса в допустимых пределах, группы крови и т.п.).
2. Временные правила определяют рамки возможных изменений значений полей элементов и признаков с привязкой к дате, событию, документу. Например, на дату записи в ЭМК у пациента не может быть определенных диагнозов, значений показателей.

Некоторые примеры правил приведены в Таблице 4.

В алгоритме расчета индекса качества данных ЭМК реализована следующая последовательность вычислений:

1. Для каждого элемента ЭМК оценивается каждый критерий качества с помощью методов, указанных в Таблице 1. Оценка происходит по логическому принципу, т.е. присваивается 0, если элемент не проходит проверку, и 1, если элемент успешно проходит проверку.
2. Для всех объектов ЭМК определяются критерии A1, A2, B1, B2, C1, C2, исходя из качества каждого элемента объекта.
3. Параметры качества A-C для ЭМК в целом рассчитываются путем усреднения параметров по каждому типу объекта ЭМК, а затем суммирования с удельным весом 50% для базовых сведений и по 25% для случаев лечения и структурированных электронных медицинских документов (СЭМД), соответственно. Значения удельных весов должны быть параметризованы и могут быть скорректированы в процессе апробации данной методики оценки качества ЭМК.

Таким образом, по каждой ЭМК формируется матрица критериев качества по типам объектов (базовые сведения, случаи лечения и СЭМДы) и по всем объектам ЭМК (таблица 5).

4. Итоговый индекс качества ЭМК рассчитывается по следующей формуле:

$$\text{Индекс качества ЭМК} = 0,25*(A1+A2) + 0,15*(B1+B2) + 0,1*(C1+C2).$$

Коэффициенты могут быть скорректированы при необходимости.

Соответственно, итоговый индекс будет иметь оценку от 0 до 100.

5. Интерпретация итогового значения индекса ЭМК должна быть откалибрована в процессе апробации методики. Для этого должны быть отобраны примеры ЭМК (до 100) с разными значениями индексов качества, а также детализированными параметрами A1, A2 и т.д. Затем медицинским экспертом либо несколькими экспертами будут оценены ЭМК на предмет клинической достоверности и полноты по 5-ти балльной шкале (недопустимое, низкое, удовлетворительное, хорошее, отличное качество). После этого будет проанализировано соответствие значения индекса качества ЭМК и мнения экспертов. Таким образом, будут получены диапазоны значений индексов ЭМК и им соответствующие мнения экспертов по 5-ти балльной шкале.

ПРИМЕР АЛГОРИТМА РАСЧЕТА

Ниже приведен практический пример расчета Индекса качества ЭМК на базе методики (Таблица 6).

Например, обязательными элементами/полями являются:

- 1) Базовые сведения в ЭМК: ФИО пациента, дата рождения пациента, ID ЭМК, дата регистрации ЭМК;
- 2) Случай: Дата открытия, дата закрытия, диагноз случая;
- 3) СЭМД: Дата документа, описание.

ЭМК, которая оценивается, имеет следующее наполнение и соответствующие параметры A1, A2 и т.д.

ОБСУЖДЕНИЕ И ЗАКЛЮЧЕНИЕ

«Индекс качества ЭМК» показывает насколько велика ценность собранной по пациенту ЭМК с точки зрения ее пригодности для машинной обработки, включая работу моделей

Таблица 4 — Примеры правил для оценки критерия правдоподобности данных

№	Элемент	Поле	Статичные правила	Временные правила
1	Социально-демографические данные	ФИО или хеш ФИО (закодированное или зашифрованное значение поля)	не должно меняться	N/A
		СНИЛС или хеш СНИЛСа	не должно меняться	N/A
		Данные о страховом полисе или хеш этих данных	не должно меняться	*
		Пол	не должно меняться	N/A
		Дата рождения	не должно меняться	N/A
		Регион проживания (прикрепления)	N/A	*
		МО прикрепления	N/A	У пациента детского возраста не может быть прикрепления к МО для взрослых
2	Минимально-необходимые клинические, антропометрические и анамнестические данные	АД	Пределы изменения АД для пациентов старше 18 лет без сопутствующих диагнозов в листе окончательных диагнозов (ЛОД)	Пределы изменения АД для пациентов старше 18 лет с сопутствующими диагнозами в ЛОД, случае
		Рост	Пределы изменения роста для пациентов старше 18 лет	N/A
		Вес	Пределы скорости изменения веса для пациентов старше 18 лет без сопутствующих диагнозов в ЛОД	Пределы скорости изменения веса для пациентов старше 18 лет с сопутствующими диагнозами в ЛОД, случае
		ЧСС	Пределы значений для пациентов старше 18 лет без сопутствующих диагнозов в ЛОД	Пределы значений для пациентов старше 18 лет с сопутствующими диагнозами в ЛОД, случае
		Общий холестерин	Пределы значений для пациентов старше 18 лет без и с сопутствующими диагнозами в ЛОД	Пределы значений для пациентов старше 18 лет с сопутствующими диагнозами в случае
		Глюкоза	Пределы значений для пациентов старше 18 лет без и с сопутствующими диагнозами в ЛОД	Пределы значений для пациентов старше 18 лет с сопутствующими диагнозами в случае
		Лист окончательных диагнозов	Проверка на наличие невозможных диагнозов исходя из возраста, пола	N/A
		Частота дыхания	Пределы значений для пациентов старше 18 лет без и с сопутствующими диагнозами в ЛОД	Пределы значений для пациентов старше 18 лет с сопутствующими диагнозами в случае
		Табакокурение	Пределы скорости изменения признака	N/A
		Креатинин	Пределы значений для пациентов старше 18 лет без и с сопутствующими диагнозами в ЛОД	Пределы значений для пациентов старше 18 лет с сопутствующими диагнозами в случае
		Гемоглобин	Пределы значений для пациентов старше 18 лет без и с сопутствующими диагнозами в ЛОД	Пределы значений для пациентов старше 18 лет с сопутствующими диагнозами в случае
		Температура	Пределы значений для пациентов старше 18 лет без и с сопутствующими диагнозами в ЛОД	Пределы значений для пациентов старше 18 лет с сопутствующими диагнозами в случае

Таблица 4 — Примеры правил для оценки критерия правдоподобности данных (продолжение)

№	Элемент	Поле	Статичные правила	Временные правила
3	Данные о случаях обращения за медицинской помощью	В ЭМК пациента должна содержаться хотя бы 1 запись о случае обращения пациента в МО, включая законченный случай, или история болезни, или случай обращения с профилактическими целями (карта диспансеризации, лист вакцинаций и т.д.). Не учитываются «Разовые посещения» – это не реальный случай обращения, а искусственный компонент цифрового профиля пациента в Webiomed.	N/A	N/A
4	Минимально-необходимые медицинские записи	Протокол инструментального исследования, в котором заполнено «Заключение»	N/A	N/A
		Протокол лабораторного исследования, в котором есть хотя бы 1 заполненный параметр (тест)	N/A	N/A
		Протокол врачебного осмотра, в котором заполнены поля «Жалобы», «Объективные данные» и указан код окончательного клинического диагноза	N/A	N/A

Примечание: * — данные правила требуют уточнения.

Таблица 5 — Расчет матрицы критериев качества для отдельных объектов ЭМК и всей ЭМК

Объект ЭМК	Параметр качества					
	A1 [0; 100]	A2[0;100]	B1[0; 100]	B2[0; 100]	C1[0; 100]	C2[0; 100]
Базовые сведения в ЭМК (вклад 50%)	0 - отсутствие хотя бы одного обязательного элемента; 100-все обязательные элементы присутствуют	100*число непустых элементов/число всех элементов в наличии	100*число технически корректных элементов/число непустых элементов	100*число корректного значения элементов/число непустых элементов	100*число правдоподобных значений элементов/число непустых элементов	100*число правдоподобных по времени элементов/число непустых элементов
Законченный случай (ЗС, вклад 25%)						
СЭМД (вклад 25%)						
ЭМК	сумма A1 всех объектов ЭМК с удельным весом 0,5 для Базовых сведений, 0,25 ЗС (сумма всех A1 у ЗС, отнормированных на количество ЗС в данной ЭМК) и 0,25 СЭМД (сумма всех A1 у СЭМД, отнормированных на количество СЭМД в данной ЭМК)	сумма A2 всех объектов ЭМК с удельным весом 0,5 для Базовых сведений, 0,25 ЗС (сумма всех A2 у ЗС, отнормированных на количество ЗС) и 0,25 СЭМД (сумма всех A2 у СЭМД, отнормированных на количество СЭМД)	сумма B1 всех объектов ЭМК с удельным весом 0,5 для Базовых сведений, 0,25 ЗС (сумма всех B1 у ЗС, отнормированных на количество ЗС) и 0,25 СЭМД (сумма всех B1 у СЭМД, отнормированных на количество СЭМД)	сумма B2 всех объектов ЭМК с удельным весом 0,5 для Базовых сведений, 0,25 ЗС (сумма всех B2 у ЗС, отнормированных на количество ЗС) и 0,25 СЭМД (сумма всех B2 у СЭМД, отнормированных на количество СЭМД)	сумма C1 всех объектов ЭМК с удельным весом 0,5 для Базовых сведений, 0,25 ЗС (сумма всех C1 у ЗС, отнормированных на количество ЗС) и 0,25 СЭМД (сумма всех C1 у СЭМД, отнормированных на количество СЭМД)	сумма C2 всех объектов ЭМК с удельным весом 0,5 для Базовых сведений, 0,25 ЗС (сумма всех C2 у ЗС, отнормированных на количество ЗС) и 0,25 СЭМД (сумма всех C2 у СЭМД, отнормированных на количество СЭМД)

машинного обучения (прогнозирования и диагностики), а также проведения исследований и разработок в сфере реальной клинической практики и создания наборов данных для машинного обучения.

Индекс имеет оценку от 0 (красная зона — ЭМК пустая и не пригодна ни для какой машинной обработки) до 100 (зеленая зона — ЭМК

имеет все минимально необходимые сведения, пригодные для анализа). Интерпретация промежуточных значений Индекса требует отдельного исследования после его внедрения с привлечением специалистов для получения «ручной оценки» набора ЭМК и сопоставления этой оценки с Индексом качества для калибровки последнего.

Таблица 6 — Пример расчета Индекса качества ЭМК

Объект ЭМК	Элемент	Пример значения элемента	A1	A2	B1	B2	C1	C2
Базовые сведения в ЭМК	ФИО пациента		0 (нет ID ЭМК)	67 (не заполнено ФИО, $100*2/3$)	50 (формат даты рождения некорректен, $100*1/2$)	100 ($100*2/2$)	50 (дата регистрации раньше даты рождения, $100*1/2$)	100
	дата рождения	06.2005						
	дата регистрации ЭМК	01.10.2000						
Случай лечения	Дата открытия	02.03.2020	100	100	100	67 (значение диагноза не соответствует МКБ-10)	100	100
	Дата закрытия	08.03.2020						
	диагноз случая	B5						
СЭМД	дата документа	06.09.2023	100	50 (не заполнено описание)	100	100	100	100
	описание	Null						
ЭМК. ИНДЕКС КАЧЕСТВА: 72,85			50	$0,5*67+0,25*100+0,25*50=71$	75	91,75	75	100

Предлагаемая методика позволяет детализировать Индекс в матричном виде по двум направлениям: критериям оценки и видам элементов ЭМК, как представлено в Таблице 6. Такая детализация позволяет системно проанализировать причины того или иного значения Индекса качества ЭМК и разработать соответствующие меры по его повышению.

Для внедрения методики необходимо разработать соответствующие словари и перечни правил, применяемых для оценки критериев. Данные словари и правила должны учитывать специфику МО, а также технические характеристики используемой МИС, для формирования ЭМК. В дальнейшем некоторые методы и алгоритмы оценки критериев могут учитываться в режиме реального времени подсвечивая врачу, какие элементы пустые, имеют недопустимые

форматы и противоречащие значения. Однако такая реализация потребует значительных работ МИС.

В данной работе были приведены примеры только базовых элементов ЭМК, однако представленная методика может быть логически расширена на любые другие элементы с учетом пожеланий заказчиков.

Представленная методика и входящий в нее алгоритм расчета индекса не являются окончательными и могут быть модифицированы в процессе эксплуатации. Также применимость и ценность данной методики должна быть подтверждена научными исследованиями. Однако она является важным шагом для осуществления интегральной и универсальной оценки качества медицинских данных, собираемых и хранимых ЭМК.

ЛИТЕРАТУРА/REFERENCES

1. Гусев А.В., Владимирский А.В., Голубев Н.А., Зарубина Т.В. Информатизация здравоохранения Российской Федерации: история и результаты развития // Национальное здравоохранение. — 2021. — №2(3). — С.5-17. [Gusev AV, Vladzimirskij AV, Golubev NA, Zarubina TV. Informatizaciya zdavoohraneniya Rossijskoj Federacii: istoriya i rezul'taty razvitiya. Nacional'noe zdavoohranenie. 2021; 2(3): 5-17. (In Russ.)] doi: 10.47093/2713-069X.2021.2.3.5-17.
2. Гусев А.В., Зингерман Б.В., Тюфилин Д.С., Зинченко В.В. Электронные медицинские карты как источник данных реальной клинической практики. Реальная клиническая практика: данные и доказательства. — 2022. — №2(2). — С.8-20. [Gusev AV, Zingerman BV, Tyufilin DS, Zinchenko V.V. Elektronnye medicinskie karty kak istochnik dannyh real'noj klinicheskoy praktiki. Real'naya klinicheskaya praktika: dannye i dokazatel'stva. 2022; 2(2): 8-20. (In Russ.)] doi: 10.37489/2782-3784-myrwd-13.
3. Johnson SG, Speedie S, Simon G, Kumar V, Westra BL. A Data Quality Ontology for the Secondary Use of EHR Data. AMIA Annu Symp Proc. 2015; 2015: 1937-46.
4. Johnson SG, Pruinelli L, Hoff A, Kumar V, Simon GJ, Steinbach M, Westra BL. A Framework for Visualizing Data Quality for Predictive Models and Clinical Quality Measures. AMIA Jt Summits Transl Sci Proc. 2019; 2019: 630-638.
5. Zhou J, Hao J, Tang M, Sun H, Wang J, Li J, Qian Q. Development of a quantitative index system for evaluating the quality of electronic medical records in disease risk intelligent prediction. BMC Med Inform Decis Mak. 2024; 24(1): 178. doi: 10.1186/s12911-024-02533-z.
6. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: Enabling reuse for clinical research. Journal of the American Medical Informatics Association, 2013. 20(1): 144-151.
7. Kahn MG, et al. A Harmonized Data Quality Assessment Terminology and Framework for the Secondary Use of Electronic Health Record Data. EGEMS (Wash DC). 2016; 4(1): 1244.
8. Кафтанов А.Н., Андрейченко А.Е., Гусев А.В. Обзор методических подходов к оценке качества ведения электронных медицинских карт // Врач и информационные технологии. — 2024. — №3. — С.6-19. [Kaftanov AN, Andrejchenko AE, Gusev AV. Obzor metodicheskikh podhodov k ocenke kachestva vedeniya elektronnyh medicinskih kart. Vrach i informacionnye tekhnologii. 2024; 3: 6-19. (In Russ.)]