

МЕТОДИЧЕСКИЕ АСПЕКТЫ СОЗДАНИЯ КРУПНОМАСШТАБНОГО КАНЦЕР-РЕГИСТРА

Ключевые слова: *канцер-регистр, информатизация, информационная технология, контроль качества данных.*

Резюме. *Рассмотрены методологические аспекты создания крупномасштабного канцер-регистра. Сформулированы основные этапы построения информационной технологии ведения канцер-регистра. Обсуждены процедуры контроля качества данных.*

За последнее десятилетие в Украине принят ряд документов относительно информатизации системы здравоохранения.

В частности для планирования и оценки противораковых мероприятий, а также проведения научных эпидемиологических и клинических исследований признано необходимым создание канцер-регистров (КР).

КР — это деятельность, направленная на сбор, хранение и обработку персонифицированной информации обо всех больных со злокачественными новообразованиями (ЗН) в пределах сферы действия регистра. Также под регистром понимается организация, занимающаяся этой деятельностью. Ценность регистра зависит от качества данных и полноты регистрации случаев заболевания.

В зависимости от сферы действия регистра можно выделить популяционные (территориальные) и больничные (госпитальные) КР. Больничные и популяционные КР являются принципиально различными системами по спектру решаемых задач.

Популяционный регистр — это организация сбора, хранения и анализа информации обо всех случаях ЗН у лиц, проживающих в пределах определенной территории (популяции). В задачи популяционного КР входит анализ заболеваемости, смертности, распространенности ЗН среди популяции определенного региона или всей страны, а также планирования развития онкологической помощи населению, организации скрининговых программ в регионе. Формирование базы данных больных со ЗН в пределах определенной территории (государства, административной единицы или определенной популяции, объединенной по какому-либо признаку, например, ликвидаторы последствий аварии на Чернобыльской АЭС) также входит в перечень задач регистра и является основой для проведения научных эпидемиологических исследований. Основной характеристикой национального популяционного КР является наличие информации обо всех случаях онкологических заболеваний, выявленных на территории Украины.

Больничный КР — это организация сбора информации о больных, поступивших в стационары онкологической клиники. В задачи больничного КР входит анализ деятельности стационаров клиники в целях ее совершенствования, анализ выживаемости больных со ЗН и оценка эффективности методов лечения. База данных о таких больных является

основой для проведения различных клинических исследований. Данные отдельных регистров могут различаться по перечню регистрируемых данных в зависимости от задач, стоящих перед регистром, и региональных потребностей. Но для международной сопоставимости данных о раке терминология и определение каждого признака должны быть унифицированы, стандартизированы подходы к регистрации заболевания. Ценность информации о зарегистрированных случаях злокачественного заболевания (ЗЗ), используемой в самых различных областях, возрастает при сохранении ее сопоставимости и преемственности с течением времени. За последние 20–25 лет развитие КР отмечено возрастающей ролью методов стандартизации, классификации и вопросов терминологии. Этот процесс возглавили Международная Ассоциация КР (МАКР, International Association of Cancer Registries, IACR), Международная Ассоциация по исследованию рака (МАИР, International Agency for Research of Cancer, IARC) в тесном сотрудничестве со Всемирной организацией здравоохранения (ВОЗ, World Health Organization, WHO). Большинство существующих КР, данные которых публикуются в монографии МАИР «Cancer Incidence in Five Continents», находятся в странах Северной Америки, Центральной и Северной Европы. Данные Восточной Европы практически отсутствуют. Отмечается насущная необходимость развития КР в странах СНГ.

До недавнего времени вся информация обо всех случаях ЗН накапливалась на бумажных носителях и практически единственным источником информации об уровнях заболеваемости на Украине служили таблицы государственной отчетности о заболеваемости и смертности от ЗН. При ручной подготовке подсчет показателей является очень трудоемким и рутинным. Кроме того, проверить качество данных и провести контроль дубликатов практически невозможно. К тому же при такой технологии отсутствует гибкость в получении новых показателей, кроме тех, которые включены в стандартные статистические таблицы. Все это приводит к недоверию и неуважительному отношению к медицинской статистике. Поэтому следующим шагом развития КР в Украине стало развитие информационных технологий сбора, хранения и обработки данных о больных онкологического профиля.

Разработка и внедрение информационных технологий (ИТ) начаты в 1989 г. в Институте онкологии АМН Украины.

Были сформулированы основные этапы построения ИТ ведения КР. На каждом этапе проектирования осуществляется тесное взаимодействие экспертов, аналитиков, прикладных программистов и будущих пользователей ИС для выработки информационных потребностей и задач.

I этап — изучение предметной области. В нашем случае — это анализ существующей системы регистрации онкологической информации: анализ принципов сбора информации; анализ всех существующих источников информации; анализ существующих в этой области медицинских документов; анализ существующих принципов сбора информации; анализ существующих информационных потоков в структуре онкологической службы и их регламентация; анализ существующих терминологических понятий и их строгое определение; анализ и классификация прикладных задач, которые планируются поставить на этих данных; проектирование иерархии задач и модель потоков данных.

II этап — изучение объекта исследования. В нашем случае — это течение онкологического заболевания: изучение признаков, которыми описывается пациент с онкологическим заболеванием, методов лечения и течения патологии; изучение особенностей используемых и существующих международных классификаций; анализ существующих подходов к контролю качества данных.

III — проектирование информационной технологии. Это поиск и определение наилучшего способа реализации прикладных систем на основе требований, выработанных на этапах анализа. На этом этапе модель предметной области преобразуется в схему баз данных и спецификации их хранения. Проектируются модели прикладных программ, работающих с данными, хранящимися в базах данных, процедуры контроля качества данных. Одновременно с этим проектируются экранные формы, форматы отчетных документов, интерфейс для взаимодействия пользователя и системы.

IV — внедрение системы. Разработка пользовательской и эксплуатационной документации. Обучение пользовательского персонала. На этом этапе осуществляется ввод данных, конвертирование файлов и проведение серии тестов на правильность работы всего программного обеспечения. Проводятся испытания на реальных данных и рабочих местах пользователей. В журнале регистрации изменений фиксируются сбои и недостатки в работе программного обеспечения, а также предложения по дальнейшему развитию технологии. Этап внедрения — один из самых ответственных, поскольку именно он определяет жизнеспособность всей системы и уровень подготовки пользовательского персонала.

Основными предпосылками разработки ИТ ведения КР стало наличие государственной регистрации больных с онкопатологией и наличие специализированной онкологической службы, представленной системой областных, районных и городских онкологических диспансеров или клиник. Принципом, положенным в основу онкологического учета, является

интеграция всей информации о больном в онкологическом учреждении по месту жительства больного. Независимо от места выявления случая заболевания или лечения пациента со ЗН в онкологический диспансер по месту жительства больного посылается соответствующий медицинский документ.

Для разработки формализованных медицинских документов был проведен анализ существующих в системе онкологического учета медицинских документов, схемы движения информации и выработаны следующие основные требования к разработке формализованных документов:

- каждый признак документа должен быть однозначно определен. Крупные онкологические форумы, съезды, конференции обычно обсуждают конкретные клинические и экспериментальные материалы, стремясь выработать методические установки и научные положения для практического использования. К сожалению, почти не обсуждаются вопросы терминологии и использования классификаций. Между тем, формулировки и терминология имеют первостепенное значение как средство общения и однозначного понимания;

- каждый признак должен иметь обоснованную функциональную нагрузку и четко определенные алгоритмы обработки;

- по возможности, не должны включаться понятия, основанные на субъективных данных и представлениях, а только объективные данные на основе результатов лабораторных исследований и клинических манипуляций или вмешательств;

- уровень детализации элементов формализованного документа должен быть одинаков;

- не должно быть функциональной зависимости между элементами документа, поскольку в противном случае необходима разработка дополнительных процедур контроля данных. Так, например, вместо регистрации возраста больного, более корректно регистрировать дату рождения и дату установления диагноза или проведения каких-либо лечебных воздействий. При этом возраст больного в любой момент можно вычислить как виртуальный признак;

- необходимо наличие организационной поддержки обеспечения репрезентативности данных, иначе высок риск получения результатов искаженных или даже не соответствующих действительности. Поэтому признаки, достоверность которых по каким-либо причинам невозможно получить, необходимо исключать из описания больного.

Исходя из изложенных выше принципов, были разработаны и утверждены МЗ Украины следующие медицинские документы: «Регистрационная карта больного злокачественным новообразованием» (далее «Регистрационная карта»); «Извещение о больном с впервые в жизни установленным диагнозом рака или с другим злокачественным новообразованием» (далее «Извещение»); «Выписка из медицинской карты стационарного больного злокачественным новообразованием» (далее «Выписка»).

Данные о смерти больных онкологического профиля регистрируются на основании данных учрежде-

ЛЕКЦИЯ

ний государственных статистических управлений. Ежемесячно проводится сопоставление данных КР с данными о смерти больных онкологического профиля, полученных на основании «Врачебных свидетельств о смерти». Указанные документы являются основными источниками информации и обеспечивают полноту регистрации информации о случаях заболевания в регионе (рис. 1).

В качестве единого регистрационного документа была определена «Регистрационная карта», которую заполняют на всех больных, которые постоянно проживают в регионе обслуживания. Все остальные документы являются информационными. Необходимо заметить, что в случае больничного КР регистрируют всех больных, поступивших в онкологическую клинику, а в случае популяционного (территориального) регистра — только тех больных, которые постоянно проживают в данном регионе, поскольку при вычислении показателей заболеваемости или смертности используются данные о населении данного региона. Каждый новый информационный документ о больном онкологического профиля дополняет информацию, имеющуюся в «Регистрационной карте». При наличии хорошо отлаженной технологии возможна замена бумажной формы хранения регистрационных документов информационной. При необходимости бумажный документ в любой момент может быть сформирован на основании информационного документа, то есть компьютерных данных. Содержащаяся в «Регистрационной карте» информация позволяет использовать данные популяционного КР не только для анализа заболеваемости и подготовки ежегодных отчетных документов, но и для дальнейшего наблюдения за больными, оценки деятельности онкологической службы, анализа выживаемости больных онкологического профиля в популяции, но и многих других приложений.

Данные популяционного КР включают в себя следующие разделы информации (см. рис. 1):

паспортные данные; данные диагноза; данные лечения; данные осмотров и патронажа; факты наблюдения за изменением состояния. В таком объеме данные со-

бираются в областных онкологических диспансерах и после контроля полноты и качества данных передаются в центральный регистр.

На этапе проектирования ИС необходимо выбрать оптимальный принцип сбора данных на уровне всей популяции, региона или государства в целом. В международной практике принят централизованный принцип сбора онкологической информации. При этом данные о больных онкологического профиля поступают непосредственно — в центральный КР. Понятно, что для такой обширной территории как Украина, с большим числом ежегодно регистрируемых больных (160 000) централизованный принцип сбора неприемлем. Поэтому для организации сбора данных на уровне Украины был выбран распределенный сбор. Информация о случаях заболевания накапливается в КР районного уровня, затем передается в областные КР и после контроля дубликатов и качества данных — в центральный КР. Каждый областной КР имеет свое территориально-административное распределение, которое поддерживается информационной технологией. Так, крупные областные КР Днепропетровской, Донецкой, Киевской, Луганской, Львовской областей имеют свои районные регистры и работают по принципу распределенного ввода. Остальные областные КР в настоящее время работают по принципу централизованного сбора данных. При наличии вычислительной техники возможно выделение новых районных веток. Таким образом, наличие вертикальных связей между центральным, областными и районными уровнями КР позволяют совместить преимущества централизованного и распределенного ведения данных: непосредственный сбор первичных данных происходит на низших уровнях, в то время как контроль качества данных, сравнительный анализ, контроль дубликатов и т.д. — наиболее высоких. Универсальное программное обеспечение и единообразие кодификаторов дает возможность инсталляции системы на любую заданную структуру районного или областного уровня и объединения данных на уровне Украины.

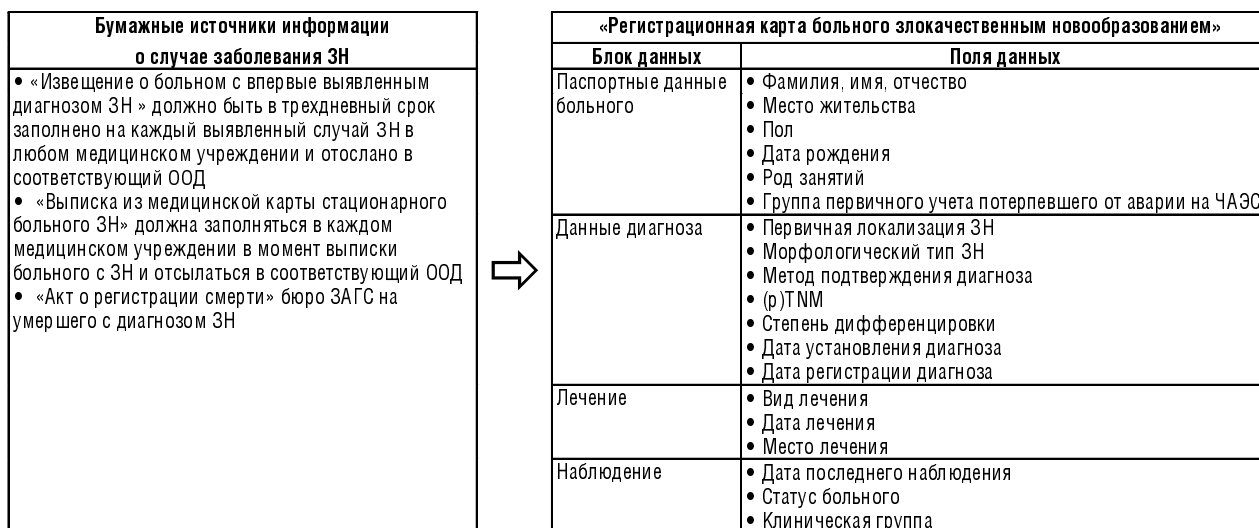


Рис. 1. Основные источники информации КР

Соблюдение международных стандартов — особенно важный момент как для международного сравнения данных, так и для внутреннего межрегионального сравнения. Проводимые по инициативе ВОЗ разработки, направленные на обеспечение сравнимости данных о заболеваемости, представлены в разработках Международной классификации болезней (МКБ) и ее специального расширения для онкологии — МКБО. МКБО позволяет регистрировать онкологические диагнозы более подробно, фиксируя для каждого новообразования как локализацию опухоли, так и ее морфологический тип и течение заболевания. Кроме того, под руководством МАИР разработана и распространена специальная программа предварительного контроля качества данных КР «IARC-CHECK». Эта программа обязательно проверяет и оценивает показатели качества данных, которые передаются для публикации в издании «Cancer in Five Continents» для обеспечения сравнимости данных на международном уровне.

Статистический код МКБ является основной единицей анализа заболеваемости и смертности от рака. Анализ состояния кодирования только этого признака в медицинских документах показал низкий уровень по сравнению с международными требованиями. По нашему мнению, причин здесь несколько: отсутствие справочной литературы по международным классификациям, в связи с чем морфологическая составляющая диагноза практически отсутствовала; субъективность кодирования и частые ошибки. Нередко кодирование статистических диагнозов поручалось лечащим врачам, которые недостаточно ознакомлены с МКБ и вообще не ознакомлены с МКБО. Поэтому к ЗН относились и пограничные, доброкачественные опухоли и даже заболевания, не имеющие отношения к новообразованиям; частая сменяемость работников локальных КР; распределенная система регистрации рака в Украине и в связи с этим многочисленность и рассредоточенность персонала КР; невозможность проведения регулярного обучения и повышения квалификации работников КР в виду финансовых трудностей. Немаловажной оказалась и психологическая причина — постоянный рост заболеваемости в «угоду» годовому отчету.

Учитывая вышеперечисленное, были разработаны автоматизированные процедуры контроля качества данных в соответствии с международными требованиями. Клинический диагноз онкопатологии хорошо поддается структурированию. В нем всегда прямым или косвенным образом можно выделить топографическую (локализация первичной опухоли) и морфологическую (гистологический тип опухоли) составляющие. Это позволяет структурно представить каждый онкологический врачебный диагноз как комбинацию конкретных значений каждой из перечисленных характеристик онкологического диагноза. Однако часто клинические диагнозы в медицинских документах содержат достаточно специфические формулировки, не включенные в алфавитный перечень МКБО, и аналоги либо обобщающие группы, кото-

рые может отыскать в МКБО только специалист. Фактически задача контроля статистических кодов сводится к автоматическому кодированию статистического диагноза по его клиническим составляющим. Для этой цели совместно с патоморфологами и клиницистами были разработаны клинические справочники-кодификаторы и установлено их функциональное соответствие статистическим международным кодам. При этом учитывалась синонимичность терминов как для топографического, так и для морфологического перечня.

Имея МКБО-составляющие для топографии, морфологии и течения заболевания, всегда можно получить код МКБ на их основе. В настоящее время в областных КР кодирование статистических диагнозов МКБ и МКБО осуществляется только автоматически. Такая технология кодирования позволяет: сохранить преемственность данных при переходе на новую версию МКБ (переход с версии МКБ-9 на МКБ-10 в 2000 г. прошел практически без проблем, поскольку переход на новую версию требует только нового файла связи между топографической и морфологической составляющими); осуществлять хранение данных более подробно, на уровне клинического диагноза, а не на уровне статистического кода диагноза; освободить от рутинного труда кодирования диагноза как работников регистра, так и лечащих врачей, что сокращает трудозатраты; улучшить качество данных; использовать разработанные клинические справочники для определения топографии и морфологии в целях ведения больничного КР.

Кроме автоматизированного кодирования статистического диагноза, осуществляется кодирование стадии заболевания по ее TNM-составляющим на основании «Международной классификации TNM», правильности регистрации случаев первично-множественного рака на основе правил регистрации, разработанных ВОЗ. Для контроля дубликатов разработана специальная технология линкиджа данных. Учитывая цель, текущие задачи КР и международные требования к качеству его данных, контроль непротиворечивости и достоверности данных должен охватывать: а) проверку наличия обязательных для регистра данных; б) контроль достоверности кодов; в) контроль достоверности и логической последовательности дат; г) контроль логической непротиворечивости данных.

Функции, которые выполняют указанные проверки, включают более чем 200 пунктов. Программа контроля качества данных постоянно модифицируется за счет расширения набора функций. Однако даже самая совершенная программа не может контролировать человеческий фактор. То есть ни один регистр не может гарантировать 100% качество данных. Но использование автоматизированных процедур контроля качества данных обеспечивает приближение к этой цели. Введение такого рода проверок требует тесного сотрудничества с патоморфологами и клиницистами, учета авторитетных мнений не одно-

ЛЕКЦИЯ

го специалиста, анализа литературных источников по организации контроля данных.

В настоящее время ИС популяционного КР внедрена во все областные онкологические центры Украины, за исключением Запорожской и Тернопольской областей. В этих областях проходит стадия внедрения. Уже сейчас у нас есть репрезентативные данные для анализа показателей в целом по Украине. Такой анализ заболеваемости и смертности от ЗН в Украине в последние 2 года проводится по данным персонализированной информации КР Украины. Автоматизированный КР позволяет перейти от цифры в статистической таблице к первичному источнику — конкретным больным, которые составили эту цифру. Поэтому статистические показатели, полученные с использованием данных регистра, являются более обоснованными и могут быть использованы как для принятия управленческих решений, так и для проведения научных исследований. Стало уже традиционным, что анализ заболеваемости и смертности населения Украины от ЗН осуществляется на основе годовых отчетных данных Министерства здравоохранения. В свою очередь эти данные основываются на информации о больных со ЗН, диагностированными и зарегистрированными в течение отчетного года. Однако вызывает сомнение возможность в течение первого года учета собрать полную информацию обо всех заболевших в отчетном году на такой обширной территории, которой является Украина. Наши исследования полноты данных показали, что в первый год регистрации в целом в Украине выявляют 94,5% от всего числа больных, во второй — 4,7% и в третий — 0,5%. Причем в отдельных областях эти цифры значительно варьируют (табл. 1). Осо-

Таблица 1
Динамика регистрации (в %) первичных больных со ЗН в КР областей Украины

Область	Год с момента установления диагноза			
	1-й	2-й	3-й	4-й
Винницкая	96,2	3,3	0,3	0,2
Волынская	95,4	3,6	0,6	0,4
Донецкая	88,3	10,4	0,9	0,4
Житомирская	96,6	2,9	0,2	0,3
Закарпатская	94,0	5,5	0,3	0,2
Луганская	97,3	1,1	0,4	0,1
Львовская	95,9	2,9	0,9	0,3
Николаевская	88,9	10,4	0,5	0,2
Хмельницкая	97,1	2,3	0,4	0,2
Черниговская	95,0	4,3	0,4	0,3
Всего	94,5	4,67	0,49	0,26

бенно значительный недоучет наблюдается в детской возрастной группе (табл. 2). Практически данные о детской заболеваемости нельзя использовать на основе только первого года учета как репрезентативные. В связи с этим анализ заболеваемости и смертности от рака можно проводить только через 2 года после полной комплектации данных. Такая практика «запаздывания» анализа наблюдается во всех КР мира. И это оправдывает цели — получение полного, качественного и объективного представления о заболеваемости и смертности населения страны от онкологических заболеваний.

Переход на автоматизированную технологию ведения данных позволяет формировать многолетнюю

базу данных о больных со ЗН в Украине. В настоящее время в базе данных национального КР содер-

Таблица 2
Динамика регистрации детей с диагнозом, установленным в 1997 г. (данные КР)

Область	Зарегистрировано (в %) от количества первичных больных, выявленных в 1997 г.		
	1997	1998	1999
Винницкая	89,2	5,4	5,4
Волынская	96,4	0,0	3,6
Луганская	98,2	1,8	0,0
Днепропетровская	84,2	10,5	5,3
Донецкая	74,4	23,3	2,2
Житомирская	91,7	0,0	8,3
Закарпатская	80,6	13,9	5,6
Львовская	90,2	7,3	2,4
Николаевская	57,7	38,5	3,8
Сумская	91,2	2,9	5,9
Харьковская	78,9	18,4	2,6
Херсонская	95,8	4,2	0,0
Хмельницкая	97,7	2,3	0,0
Черниговская	94,7	5,3	0,0
Всего	86,5	10,7	2,6

жится информация о более 1,5 млн таких больных. Такая база данных является ценным материалом для проведения различного рода эпидемиологических исследований, анализа выживаемости популяции больных онкологического профиля. Особенно важны такие исследования для оценки возможного влияния последствий аварии на Чернобыльской АЭС.

ЛИТЕРАТУРА

1. Cancer Epidemiology: Principles and Methods. (ed. by Isabel dos Santos Silva). Lyon: IACR, 1998. 442 p.
2. Survey of Cancer Registries in the European Union. (ed. by H Storm, L Clemmensen, R Beack). Technical Report № 28. Lyon: IACR, 1998. 64 p.
3. Cancer registration principles and methods. (ed. by OM Jensen, DM Parkin, R Maclellan, et al). Scientific Publication № 95. Lyon: IACR, 1991. 285 p.
4. Automated Data Collection in Cancer Registration (ed. by R Beack, L Simonato, H Storm, E Demaret). Technical Report № 32. Lyon: IACR, 1998. 52 p.
5. Comparability and Quality Control in Cancer Registration. (ed. by M Parkin, V Chen, J Ferlay, et al) Technical Report № 19. Lyon: IACR, 1994. 110 p.
6. Classification International of Diseases for Oncology. Second Edition. World Health Organisation, Geneva, 1990. 144 p.
7. Multiple Primaries. (IARC/IACR Internal Report № 94/003). Lyon: IACR, 1994.

METHODOLOGICAL ASPECTS OF DEVELOPING A LARGE-SCALE CANCER REGISTRY

L.O. Goulak

Summary. *The paper considers methodological aspects of developing a large-scale cancer registry. The basic phases of designing an informational technology for maintaining a cancer registry are formulated. The procedures for data quality control are discussed.*

Key Words: cancer registry, informatization, information technology, data quality control, malignant neoplasm.