

Контроль качества данных популяционного ракового регистра

Популяционные раковые регистры должны быть в состоянии предоставить объективные индикаторы качества собранных ими данных. Описание доступных методов было дано в одном из первых технических отчетов МАИР (Parkin *et al.*, 1994). В 2009 г. их обновленный вариант был представлен в двух статьях (Parkin и Bray, 2009; Bray и Parkin, 2009). В этих работах авторы описывают четыре аспекта качества данных: сопоставимость, достоверность, своевременность и полнота.

1. Сопоставимость

Сопоставимость статистических данных, относящихся к разным популяциям и разным временным периодам в пределах одной популяции, обеспечивается путем стандартизации процедур классификации и кодирования новых случаев ЗНО, а

также использования согласованных подходов к их учету, например, таких, как правила регистрации первично-множественных злокачественных новообразований у одного больного. Стандартом классификации и кодирования ЗНО является опубликованная ВОЗ система кодирования МКБ-О, предоставляющая стандарты кодирования топографии (локализация опухоли в организме), морфологии (микроскопический тип опухоли), поведения опухоли (злокачественная, доброкачественная или карцинома *in situ*), а также степени ее дифференцировки. МКБ-О-3 также обеспечивает стандартную схему кодирования метода установления диагноза и правила МАИР по кодированию первично-множественных ЗНО. Канцерогенез является процессом, который иногда может продолжаться десятилетия, поэтому определение даты начала заболе-

вания остается достаточно произвольным. Особое значение приобретает соблюдение согласованных стандартов в этом вопросе. Правила определения даты установления диагноза были предложены Европейской сетью раковых регистров (<http://www.enccr.eu/images/docs/recommendations/recommendations.pdf>).

В странах с низким и средним уровнем дохода соблюдение международных стандартов может быть сопряжено с трудностями объективного характера. Например, недостаточный охват случаев ЗНО патоморфологической службой или трудности с доступом к диагностическим службам приведут к снижению процента морфологически верифицированных случаев. В свою очередь, это станет причиной задержки в определении даты установления диагноза, поскольку Европейская

сеть раковых регистров рекомендует использовать в качестве последней дату первого гистологического или цитологического подтверждения наличия злокачественных клеток.

2. Достоверность

Точность регистрируемой информации значительно повышается, если при вводе данных одновременно проводятся проверки на их согласованность и внутреннюю непротиворечивость. Примером может служить набор проверок, включенных в систему CanReg (см. Приложение 1). Большинство регистров также проводят формальные или неформальные проверки аккуратности выполняемой сотрудниками работы, например, пересматривая информацию от одного или нескольких источников данных с целью проверки точности зарегистрированных данных или организуя соответствующие занятия и упражнения. При выявлении недостатков в работе сотрудников принимаются меры по их исправлению.

Большинство регистров предоставляют информацию о трех показателях, характеризующих точность регистрируемой информации. К ним относятся:

- доля (или процент) случаев заболевания с отсутствующими данными
- доля больных с морфологически верифицированным диагнозом (МВ%)
- доля случаев заболевания, зарегистрированных только на основании свидетельства о смерти (ТСС%).

2.1 Доля (или процент) случаев заболевания с отсутствующими данными

Доля случаев с неуказанными или неизвестными значениями некоторых переменных, таких как возраст или стадия ЗНО, также является по-

казателем качества данных. Важным показателем качества данных является доля случаев ЗНО с неопределенной локализацией первичной опухоли (англ. primary site uncertain, PSU%). В дополнение к коду МКБ-О для неуточненной локализации ЗНО (C80.9), эта категория должна также включать другие неточно обозначенные локализации ЗНО.

Сбор некоторой информации в странах с низким и средним уровнем дохода может быть сопряжен с большими трудностями. Например, сказанное в полной мере относится к получению и последующему использованию общенационального персонального идентификационного номера пациента, что в будущем потребует применения более сложных и менее точных процедур соединения информации из разных источников данных. Многие СНСД испытывают трудности либо с получением данных о смертности, либо с их низким качеством. Это может создавать многочисленные проблемы для ракового регистра, например, такие, как неполная регистрация ЗНО из-за отсутствия данных о случаях рака, зарегистрированных только на основании свидетельств о смерти (ТСС), и неспособность вычислить стандартные показатели качества данных (за исключением доли случаев с морфологически верифицированным диагнозом [МВ%]). При отсутствии данных о смертности единственным путем оценки полноты регистрации ЗНО является использование метода независимого выявления случаев заболевания или метода двойного охвата (их описание приводится ниже).

2.2 Доля больных с морфологически верифицированным диагнозом (МВ%)

Морфологическая верификация (морфологическое подтверждение)

диагноза означает, что диагноз рака был установлен по результатам гистологического или цитологического исследования. Этот метод традиционно считается своего рода «золотым стандартом», как бы подразумевающим, что диагностика злокачественных новообразований с помощью других методов является менее точной. Однако вызывает сомнение справедливость утверждения, что результаты эксфолиативной цитологии всегда бывают более точными, чем результаты МРТ или КТ. Как правило, считается, что высокое значение МВ% характеризует точность диагностики, а низкое – недостаточную достоверность информации.

При рассмотрении данных ПРР, представленных для публикации в издании «Заболеваемость раком на пяти континентах» (англ. C15) проводятся т.н. редакционные проверки. Они включают формальное сравнение значений МВ% (в разбивке по полу и основным локализациям) со «стандартным уровнем», основанным на значениях, наблюдаемых в этом же регионе 5 лет назад. Приложение 2 содержит таблицы со «стандартными» значениями избранных показателей качества данных (включая МВ%) в разрезе стран или регионов, применяемые в процессе подготовки публикации C15. Если значения МВ% статистически значимо ниже ожидаемого уровня, могут возникнуть сомнения в достоверности данных. Следует помнить, что раковые регистры обычно не могут повлиять на наличие или использование патоморфологических услуг в охватываемом регионе. С другой стороны, в странах с низким и средним уровнем дохода наблюдается противоположная ситуация, характеризующаяся достаточно высокими значениями МВ%, что также является поводом для беспокойства. Сбор данных о слу-

чаях рака путем обращения в патоморфологические отделения – это значительно более легкая процедура, чем поиск соответствующих данных в клинических отделениях или в плохо организованных архивах больниц. Значительная доля ЗНО, выявленных через патоморфологические отделения, может указывать на дефекты в системе выявления ЗНО и, следовательно, на неполноту их регистрацию. Более того, неполнота данных будет усугубляться из-за отсутствия информации о ЗНО, диагноз которых трудно поставить с помощью биопсии (например, заболевания легких, печени, головного мозга и поджелудочной железы).

2.3 Доля случаев, зарегистрированных только на основании свидетельства о смерти (ТСС%, DCO%)

В эту категорию входят зарегистрированные случаи ЗНО, единственным источником информации о которых является свидетельство о смерти. Как было отмечено ранее, характер свидетельств о смерти в СНСД широко варьируется – от свидетельств, выданных отделом регистрации актов гражданского состояния, до свидетельств, выданных моргом больницы. Тем не менее, почти во всех случаях точность информации о диагнозе и причинах смерти является сомнительной, поскольку сотрудник, заполняющий свидетельство о смерти, вряд ли мог иметь достаточное число контактов с пациентом перед смертью и/или может иметь недостаточную подготовку для заполнения свидетельства о смерти. Поэтому при отсутствии медицинских свидетельств о наличии при жизни онкологического заболевания у лица, предположительно умершего от рака или страдающего этим заболеванием на момент смерти,

имеется вполне обоснованное подозрение о неправильности диагноза. Тем не менее, практика ведения регистра требует включения таких случаев ЗНО. Если же они составляют значительную часть от общего числа зарегистрированных злокачественных новообразований, достоверность данных может вызывать сомнения.

Установление объективных критериев приемлемой доли случаев ТСС является трудной задачей, поскольку ее величина зависит от таких факторов, как наличие и доступность свидетельств о смерти, успешность установления взаимосвязей между данными из внешних источников информации и базой данных регистра, качество заполнения причин смерти, а также технические возможности для выяснения, не был ли ТСС-случай ЗНО уже зарегистрирован где-либо ранее.

2.4 Внутренняя согласованность

Перед сохранением вновь введенных сведений в центральной базе данных их следует проверить на достоверность, внутреннюю согласованность и согласованность с уже имеющимися данными регистра. Такую проверку следует проводить после каждого изменения данных.

3. Своевременность

Часто от раковых регистров требуют предоставления отчетности в максимально короткие сроки. Тем не менее, и сотрудникам раковых регистров, и потребителям информации, следует осознавать необходимость компромисса между своевременностью представления данных и их полнотой. Своевременность зависит от скорости, с которой регистр способен собирать,

обрабатывать и предоставлять в достаточной степени полную и точную информацию. В некоторых странах, например, в Соединенном Королевстве, электронные средства сбора данных позволили ускорить процесс регистрации случаев рака. Некоторые сети раковых регистров, такие как SEER и Североамериканская ассоциация центральных раковых регистров, имеют соглашение, согласно которому их члены должны предоставлять данные в течение 22-24 месяцев после окончания года установления диагноза. Регистры могут применять также математические методы, такие, как модель с запаздыванием (в рамках которой осуществляется оценка недоучета случаев заболевания на момент публикации отчета) или осуществлять построение краткосрочных прогнозов с целью получения оценки числа случаев рака в текущем году.

4. Полнота данных

Parkin и Bray (2009) проводят различие между

- качественными (или полуколичественными) методами, которые дают представление о степени полноты данных в сравнении с другими регистрами или периодами времени, и
- количественными методами, позволяющими получить численную оценку степени полноты уже зарегистрированного числа случаев рака.

4.1 Полуколичественные методы

Из числа существующих полуколичественных методов ранее уже отмечалась возможность анализа значений доли морфологически верифицированных случаев: относительно высокие значения пока-

зателя МВ% могут свидетельствовать о неполном сборе данных.

Случай злокачественного новообразования может быть выявлен на основании данных из различных источников информации: лечебных учреждений, лабораторий, свидетельства о смерти. Большое значение числа источников (извещений) в перерасчете на один выявленный случай рака обычно свидетельствует о том, что «нулевые» источники (т.е. не выявившие ни одного случая) будут довольно редким явлением. Другими широко используемыми показателями являются:

- индекс достоверности учета (т.е. отношение числа умерших от ЗНО к числу заболевших ЗНО за отчетный период времени)
- стабильность показателей заболеваемости во времени
- сопоставление показателей заболеваемости с соответствующими показателями в других (подобных) популяциях.

4.1.1 Индекс достоверности учета (отношение числа умерших от ЗНО к числу заболевших ЗНО за отчетный период времени)

Индекс достоверности учета (ИДУ) является важным показателем, который широко используется – например, в СИ5 – для выявления возможной неполноты данных. ИДУ представляет собой отношение количества смертей от определенного вида ЗНО, полученного от независимого от регистра источника данных (как правило из системы регистрации актов гражданского состояния), к количеству новых случаев этого же вида ЗНО, зарегистрированных в тот же период времени. Однако для использования этого метода необходимы высококачественные данные смертности (в частности, регистрации точной при-

чины смерти), поскольку тогда ИДУ приблизительно равен величине: 1 минус показатель 5-летней выживаемости. Это позволяет определить объективные стандарты значений ИДУ, которые являются применимыми к регионам с более или менее аналогичными показателями выживаемости (см. Приложение 2). Данный метод не может быть использован для стран и регионов без эффективной системы регистрации смертности, а также в том случае, когда причина смерти неизвестна или может быть неточной – ситуация, характерная почти для всех стран Африки и многих стран Азии.

4.1.2 Стабильность показателей заболеваемости во времени

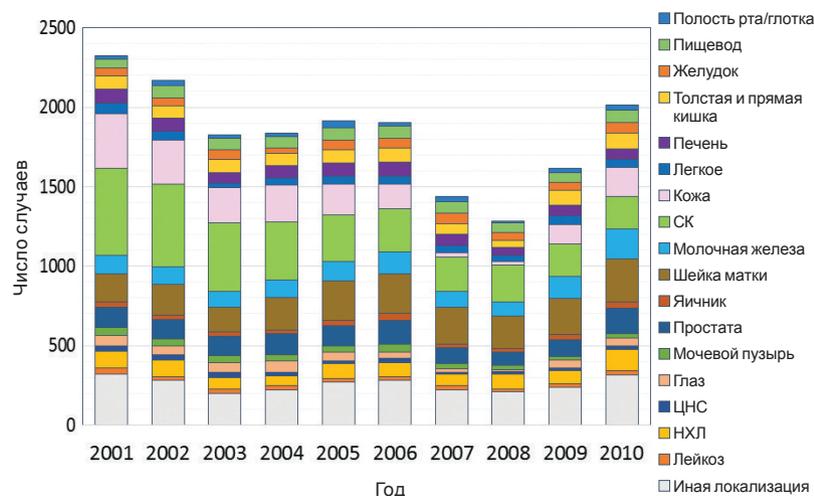
Любой раковый регистр может осуществить проверку числа ежегодно регистрируемых злокачественных новообразований. При отсутствии явно выраженных изменений в численности населения это поможет быстро определить потенциальные дефекты в системе выявления случаев рака.

На рис. 5.1 приведен соответствующий пример. Наблюдается очевидная нехватка информации о случаях рака за 2007, 2008 и 2009 годы по большинству локализаций, но в особенности – по раку кожи.

4.1.3 Сравнение показателей заболеваемости между (подобными) популяциями

Конечно же, не все популяции будут иметь одинаковую структуру показателей заболеваемости. Поэтому мониторинг их различий является одной из задач ракового регистра. Тем не менее, с целью выявления возможных различий желательно сравнивать результаты исследуемого регистра с результатами регистров, охватывающих аналогичную группу населения (по географическому признаку или по этническому составу), при условии, что все рассматриваемые регистры предоставляют высококачественные данные. Некоторые отличия вполне ожидаемы, или они могут быть объяснены воздействием известных факторов риска. Однако систематические различия

Рисунок 5.1. Количество новых случаев заболевания раком в разрезе локализаций (данные ракового регистра за 2001-2010 гг.). ЦНС – центральная нервная система; СК – саркома Капоши; НХЛ – неходжкинская лимфома.



в показателях, особенно в случае, когда большинство наблюдаемых значений показателей ниже ожидаемых значений, могут вызывать подозрения о неполной регистрации случаев рака.

Этот метод используется при подготовке публикации C15, когда для каждого регистра его данные сопоставляются с данными группы регистров из той же страны (или географической зоны) (Приложение 2).

4.2 Количественные методы

Для количественной оценки полноты регистрации данных используются три метода:

- метод независимого выявления ЗНО [англ. independent case ascertainment]
- метод двойного охвата [англ. capture–recapture method]
- метод, основанный на анализе свидетельств о смерти [англ. death certificate methods].

4.2.1 Независимое выявление злокачественных новообразований

Сравнение данных ракового регистра с базами данных, содержащими случаи ЗНО, но собранными независимо от процедур выявления случаев рака канцер-регистром, является особенно полезным и объективным методом оценки полноты данных. Необходимым условием использования этого метода является возможность проведения автоматизированного или ручного связывания информации между базой данных ракового регистра и базой данных о злокачественных новообразованиях, выявленных независимо от регистра. В этом случае возможно определить число случаев ЗНО, которые не попали в регистр. Тогда доля онкологических пациентов, уже зарегистрированных в канцер-ре-

гистре, будет прямой количественной оценкой полноты данных.

Существование таких бумажных списков (или электронных файлов) больных раком, проживающих на территории действия ракового регистра – например, созданных при проведении научных исследований или опросов – предоставляет хорошую возможность для оценки полноты данных ПРР, которую не следует упускать.

4.2.2 Метод двойного охвата

Аналогично расчету числа извещений на один случай заболевания раком, в этом методе также используется тот факт, что раковый регистр получает уведомления об одном и том же случае ЗНО из нескольких источников. Обычно при использовании этого метода источники данных группируются по категориям: в одной находятся все лечебные учреждения (и лаборатории), в другой – свидетельства о смерти; эти категории можно считать более или менее независимыми друг от друга. Основная идея заключается в том, что зная число случаев, зарегистрированных только одним, какими-то двумя или всеми тремя источниками, можно оценить число случаев, не зарегистрированных ни одним, т.е. число пропущенных случаев. На практике анализ полноты данных с использованием метода двойного охвата требует успешного связывания сведений об источниках информации, известивших об одном и том же случае ЗНО. Для пользователей программного продукта CanReg5 это не является проблемой, так как в этой системе источники информации о каждом случае заболевания раком хранятся в одной таблице, связанной с данными о диагнозе. Благодаря такой структуре файлов анализ данных проводится особенно просто.

4.2.3 Метод оценки полноты выявления злокачественных новообразований на основании свидетельства о смерти

Для использования этого метода необходимо, чтобы на территории действия ракового регистра существовала возможность определения и регистрации причины смерти с максимальной точностью и полнотой. Поэтому во многих СНСД применение этого метода будет затруднено. С другой стороны, предшествующие два метода могут применяться во всех СНСД.

5. Показатели качества данных для расчета популяционного выживаемости онкологических больных

В отличие от оценки уровня заболеваемости, для расчета оценок популяционной выживаемости необходимы высококачественные данные о жизненном статусе больных, состоящих на диспансерном учете [т.н. прослеженность]. В идеальном случае для получения информации о прослеженности больных необходимо, чтобы регистр имел доступ к данным о смертности (на уровне причины смерти) и чтобы были в наличии эффективные процедуры связывания информации между базой данных регистра и базой данных смертности. В последнем случае наилучшим решением будет использование уникального, возможно, общенационального, идентификационного номера. Поскольку системы регистрации актов гражданского состояния в СНСД часто либо отсутствуют, либо ненадежны или недоступны для использования раковыми регистрами, в СНСД интенсивно используются методы активного наблюдения пациентов. Поскольку ТСС-случаи и случаи рака у больных, для которых полностью отсутствует информация о после-

дующем наблюдении, исключаются из расчета показателей выживаемости, их частоты являются индикаторами качества данных для расчета популяционной выживаемости (Swaminathan *et al.*, 2011). Выбытие (потеря) больного из-под наблюдения является причиной появления систематической ошибки в показателях выживаемости, что имеет место в том числе в регистрах стран с высоким уровнем дохода, поскольку даже небольшая недооценка количества умерших может привести к переоценке показателей долгосрочной выживаемости (Brenner и Nakulinen, 2009). В СНСД, характе-

ризующихся плохо функционирующей системой сбора медицинской статистики и отсутствием данных о смертности, оценки показателей онкологической выживаемости, полученные от ПРР, иногда являются единственной возможностью изучения состояния онкологической помощи в стране.

Как отметил Skeet в публикации *Cancer Registration: Principles and Methods* [Регистрация рака: Принципы и методы] (Skeet, 1991), «все регистры должны быть в состоянии указать некоторый объективный критерий [оценки качества], а не полагаться на полученную

мудрость и благие надежды.» Это хороший совет, к которому, однако, не всегда прислушиваются. При составлении отчетов о результатах работы регистра необходимо предварительно провести оценку их качества. Полученные результаты будут почти несомненно использоваться для сравнения показателей онкологической заболеваемости или онкологического риска между подгруппами населения и временными периодами, поэтому важно удостовериться, что полученные результаты не искажены теми или иными дефектами процесса регистрации случаев заболевания.

Ключевые положения

- Все популяционные раковые регистры должны быть в состоянии обеспечить определенную степень объективного контроля качества собранных данных.
- Доступные методы должны быть описаны и обновлены, и они должны охватывать четыре основные составляющие качества данных: сопоставимость, достоверность, своевременность и полнота.