

Объем выборки для популяционного изучения общей урологической заболеваемости

Sample size for population study of total urological incidence

F.A. Akilov, B.M. Mamatkulov, U.A. Khudaybergenov

The paper is dedicated to methods of determination of sample size for population incidence study.

An analysis of incidence including urological incidence is necessary condition for achievement of high level population health, for intensive health care service development, for control of medical and preventive treatment facility function and management thereof. Determination of true prevalence rate of urological diseases needs for evaluation of so called hidden urological prevalence. To resolve this task focused population examination is more informative contrary to analysis of medical aid appealability.

However especial population examinations in general are labor-consuming and associated with significant time and resources expenses. In particular it concerns total survey. For this reason sampling examinations as an optimal variant of true prevalence evaluation are used frequently.

One of the essential constituent and at the same time one of stages of sampling population examination undertaken for incidence determination is sample volume (size) estimation. Correct estimation of necessary case number for sampling incidence examination performing substantially provides objectivity of the data obtained and determines requirement for facilities and resources for its realization.

An equation is provided for sample size computing in case of simple random selection wherein frequency of studied attribute in population (for example, urological incidence) and frequency of alternative attribute are considered. Literature review about urological prevalence study and methods of sample size determination is presented.

Ф.А. Акилов, Б.М. Маматкулов, У.А. Худайбергенев

Республиканский специализированный центр урологии, Ташкент, Узбекистан

Анализ показателей заболеваемости, в т.ч. урологической, является необходимым условием достижения высокого уровня здоровья населения, интенсивного развития здравоохранения, контроля за деятельностью лечебно-профилактических учреждений и управления ими [1].

Урологические заболевания оказывают существенное влияние на состояние здоровья общества [2, 3]. Если в 80-90-х гг. XX столетия доля урологических болезней составляла 4-5% общей заболеваемости, то к нынешнему времени она увеличилась до 7%. В действительности этот показатель еще выше (предположительно около 9%), т.к. более 30% урологической патологии в соответствии с международной классификацией болезней (10-го пересмотра) относится к другим классам заболеваний [4].

По данным обращаемости населения за медицинской помощью в 2006 г. в России было зарегистрировано 14 909 651 случаев болезней мочеполовой системы [1]. В Узбекистане в том же году было зафиксировано 4900,60 за-

болеваний органов мочеполовой системы на 100 000 населения. При этом число впервые обнаруженных заболеваний мочеполовой системы в 2006 г. составило 2610,89 на 100 000 населения Республики [5].

Определение истинной распространенности урологических болезней требует оценки т.н. скрытой урологической заболеваемости («урологический айсберг») [6]. Для решения этой задачи более информативно целенаправленное обследование населения, а не анализ данных его обращаемости за медицинской помощью [7]. Это связано с тем, что обращаемость зависит от множества факторов, таких как развитие инфраструктуры и медицинской помощи, ее доступность, общая и медицинская культура, привычки, обычаи и менталитет населения, влияющие на его медицинскую активность [8].

Однако специальные обследования населения обычно трудоемки и связаны со значительными затратами времени и ресурсов. В особенности это касается сплошного обследования. Поэтому в качестве оптимального варианта

изучения истинной заболеваемости нередко используют выборочные обследования.

Теория и практика выборочного метода обследования хорошо разработаны [9, 10, 11]. С другой стороны, имеются указания на то, что рекомендации, содержащиеся в некоторых руководствах, слишком сложны, а основанные на них практические предложения не всегда подтверждаются и применимы в реальных условиях [12].

Одной из неотъемлемых составляющих и одновременно одним из этапов выборочного популяционного исследования, предпринимаемого с целью определения заболеваемости, является расчет объема (размера) выборки. Правильное определение необходимого числа наблюдений при проведении выборочного изучения заболеваемости в значительной мере обуславливает объективность полученных данных и определяет потребность в силах и средствах для его осуществления [13].

На практике объем выборки определяют по формулам [9, 10, 11, 14], номограммам [15], таблицам [13, 16], а также посредством компьютерных программ [17, 18].

Среди нескольких способов отбора населения для обследования, т.е. формирования выборки, самой элементарной является т.н. простая случайная выборка (простой рандомизированный отбор) [19]. При этом любая единица выборки имеет равные шансы быть отобранной с помощью жеребьевки, таблиц или компьютерного генератора случайных чисел. Подчеркивается, что произвольный или беспорядочный (бессистемный) отбор не попадает под определение случайного отбора [10] и, следовательно, его применение не соответствует правилам выборочного метода.

Отмечается, что необходимым условием для осуществления любого способа отбора является предварительное составление основы выборки – поименного списка лиц, из которого и будет проводиться отбор [20]. Указанная основа выборки составляет т.н. генеральную совокупность, или популяцию.

При простом случайном отборе объем выборки, или выборочной совокупности, может быть определен на основании следующей формулы¹ [21]:

$$n = \frac{pq}{(E/1,96)^2} \quad (1)$$

где n – численность выборочной совокупности (объем, размер выборки),
 p – частота изучаемого признака в популяции (например, урологической заболеваемости), %,
 q – частота альтернативного признака, т.е. его отсутствия ($100 - p$), %,
 E – ошибка выборки.

Коэффициент 1,96 в данной формуле свидетельствует о принятом 95%-ном доверительном интервале оценки показателя [19], т.е. о том, что предполагаемая величина заболеваемости будет определена с 95%-ной вероятностью.

Кроме того, для определения объема выборки предложены формулы, принимающие в расчет численность общей популяции, из которой производится выборка [14, 19, 22].

Объем выборки зависит от распространенности изучаемого явления (в данном случае – урологических заболеваний), допускаемой точности исчисляемых показателей, а также от особенностей организации исследования [13]. Таким образом, для определения объема выборки изначально необходимо знать показатель распространенности урологических заболеваний. Очевидно, что при такой постановке вопроса проблема является собой парадокс, так

¹ Представлена одна из наиболее простых, доступных для понимания и наглядных формул.

как далеко не всегда исходный уровень заболеваемости известен, ведь именно его и требуется в конечном итоге определить. Решение, казалось бы, патовой ситуации может быть достигнуто одним из трех способов [14]: получением ориентировочных данных при помощи предварительного пробного исследования; использованием литературных источников; принятием максимального возможного значения распространенности и его альтернативного параметра, т.е. 0,5 или 50%. Впрочем, отмечается, что в последнем случае численность выборки будет неоправданно завышена, и использовать его приходится лишь при отсутствии данных пробного исследования или сведений из литературы.

При вычислении объема выборки для оценки распространенности в популяции всех урологических заболеваний (общей урологической заболеваемости) исходная информация может быть добыта из литературных источников.

По сведениям Л.А. Сковердяк (1975), распространенность заболеваний почек и мочевыводящих путей по данным обращаемости и комплексных медицинских осмотров равна 2,02% [23]. Однако специальные осмотры позволили выявить автору дополнительно 4,5% пациентов с урологическими болезнями. Истинная же урологическая заболеваемость может превосходить соответствующий показатель по обращаемости в несколько раз [24, 25]. Кроме того, при анализе литературы выяснилось, что урологическая заболеваемость по отдельным нозологическим единицам, к примеру мочекаменной болезнью, значительно превышает приведенный Л.А. Сковердяк показатель общей заболеваемости [26]. Исходя из этого, а также ввиду более чем 30-летней

давности сведений Л.А. Сковердяк, относиться к ним как к основе для расчета выборки следует критически.

По данным Б.С. Кожобекова (1998), в течение 2,5 лет было обследовано 5323 человека, проживающих в Казалинском и Аральском районах Приаралья (Казахстан) [27]. Посредством ультразвукового исследования у 3187 из них (59,9%) была обнаружена патология мочевыводящей системы: хронический пиелонефрит – 1507 человек, нефроптоз – 719, мочекаменный диатез – 660, мочекаменная болезнь – 53, киста почки – 63, вторично сморщенная почка – 27, пиелозатекание – 34, острый пиелонефрит – 20, аномалии – 25, гломерулонефрит – 36, гидронефроз – 11 и прочие заболевания – 32 пациента. При анализе этих данных следует отметить, что возможность диагностики некоторых состояний (таких, как мочекаменный диатез, пиелонефрит, гломерулонефрит) только лишь при помощи УЗИ нам представляется спорной. Исходя из этого, можно предположить, что распространенность обнаруженных патологических изменений мочеполовой системы составила в данном исследовании 18,1% (964 из 5323 обследованных).

В трех отобранных регионах Боливии, представляющих различные географические и социальноэкономические условия – тропическом, равнинном и долинном – в течение более чем 7 месяцев были вовлечены в исследование 14 082 предположительно здоровых человека [28]. Средний возраст обследуемых составил 20,3 года, мужчин было 48% и женщин – 52%. Пациенты были подвергнуты скринингу с использованием тест-полосок и микроскопии осадка мочи. У 4261 (30,3%) из них были обнаружены изменения в моче: гематурия

(14,3% обследуемых), лейкоцитурия (12,5%), протеинурия (2,1%), протеинурия в сочетании с глюкозурией (0,3%), а также протеинурия в сочетании с гематурией (0,1%). Обследование 1019 пациентов из этой группы удалось продолжить вплоть до установления окончательного диагноза, что потребовало выполнения биохимического и микробиологических исследований, УЗИ и в необходимых случаях внутривенной пиелографии. В итоге у 324 (48,4%) была диагностирована инфекция мочевого тракта, у 294 (43,9%) – изолированная гематурия, у 11 (1,6%) – хроническая почечная недостаточность, еще у 11 (1,6%) – туберкулез почки и у 29 (4,3%) – смешанная патология (камни почек – 1,3%, диабетическая нефропатия – 1%, поликистоз почек – 1,9%).

В украинском исследовании [29] в различных группах организованной популяции обследованы 2620 человек, в т.ч. 1256 мужчин и 1364 женщины. Средний возраст обследованных мужчин составил 44,3 года, женщин – 43,6 года. Урологические заболевания выявлены у 11,1% мужчин и 10,7% женщин, причем в младшей возрастной группе ни у мужчин, ни у женщин эти болезни не встречались.

При изучении общественного здоровья Бостонского региона (Массачусетс, США) обследовано 5506 взрослых жителей Бостона обоего пола в возрасте от 30 до 79 лет [30]. Случайная выборка была сформирована на основании расслоенного двухстадийного кластерного дизайна. Сведения о наличии симптомов нижних мочевых путей (СНМП) были получены путем интервьюирования и с использованием опросников. Распространенность СНМП в изученной популяции составляла 18,7%, значительно повышаясь с воз-

растом, однако не была связана с полом или этнической / расовой принадлежностью. Недостатком данного исследования является неполный охват возрастных групп населения, а также ориентация только на наличие симптомов, выявленных с помощью опросников. Между тем очевидно, что далеко не всегда и не все урологические заболевания проявляются выраженной симптоматикой [28], что могло привести к недостаточному и неполному их выявлению.

Тем не менее, из всех вышеупомянутых исследований бостонское выделяется большей тщательностью планирования и подробным описанием. Поэтому показатель распространенности СНМП, определенный в его итоге, вероятно, более приемлем для подстановки в формулу с целью определения объема выборки. Однако выбор и предпочтение исходного показателя является прерогативой исследователя.

Что же касается ошибки выборки (Е), то ее предварительная оценка (предпочитаемая величина перед подстановкой в формулу) часто произвольна и остается на усмотрении исследователя. Как правило, не рекомендуется принимать ошибку выборки выше 5% (в абсолютном значении²) [21]. Согласно А.В. Чаклину и соавт. (1990), величина ошибки выборки не должна превышать 25% от величины предполагаемого показателя распространенности [22].

Приведем пример расчета минимального объема простой случайной выборки, необходимого для оценки урологической заболеваемости.

Воспользуемся для расчета

² К примеру, ожидаемая распространенность урологических заболеваний принимается равной $20 \pm 5\%$, где 5% – допускаемая исследователем ошибка выборки. В итоге исследователь мирится с тем, что фактически определяемый им показатель распространенности будет находиться в промежутке от 15 до 25%.

значением потенциального показателя урологической заболеваемости, определенной бостонским исследованием (18,7%) и возможной ошибки выборки 5%. Подставив эти значения в формулу (1), имеем:

$$n = \frac{18,7 \times 81,3}{(5/1,96)^2} = \frac{1520,31}{6,51} = 233,5$$

Таким образом, для оценки урологической заболеваемости при заданных параметрах ее вероятной величины, базирующейся на предпочтенных сведениях литературы, потребуется отобрать 234 человека из взятой популяции (генеральной совокупности).

Существуют и более сложные способы формирования выборочной совокупности, отличающиеся от простой случайной выборки (стратифицированный, кластерный, многоступенчатый, комбинированный).

Так, при стратифицированном³, или расслоенном, способе отбора все население разделяют на классы, именуемые слоями (стратами), в зависимости от изучаемых характеристик, таких как возраст, пол и т.п., после чего из каждого слоя отбирается простая случайная выборка с одинаковой или специально рассчитанной (для каждого слоя) выборочной долей [19, 20]. Тем самым достигается лучшая представительность возрастнo-половой структуры населения, попавшего в выборку, по отношению ко всему населению (генеральной совокупности).

При кластерном⁴, или гнездовом, отборе совокупность предварительно разделяют на группы (кластеры) из однородных единиц (например, по географическому принципу), далее выполняют случайную выборку кластеров, и в выбранных кластерах прово-

дят сплошное обследование всех единиц [11, 19]. Примером может служить обследование всех жителей нескольких отобранных деревень в сельском районе.

При использовании выше-названных сложных способов выборочного исследования объем выборки, рассчитанный для простого случайного отбора, оказывается недостаточным, поскольку в этом случае вступает в действие т.н. дизайн-эффект (design effect, DEFF) [12]. Последний представлен в виде умножающего коэффициента, который рассчитывается для каждого типа исследования [12, 16].

При необходимости сравнительного анализа заболеваемости в различных группах населения (социально-экономических, профессиональных и др.) и для изучения распространенности редких заболеваний требуемый объем выборки должен быть увеличен и соответствующим образом рассчитан [13].

В свете этого примечательно, что для упомянутого бостонского исследования объем выборки составил 5506 человек, несмотря на то что, согласно общеизвестным формулам и таблицам, предполагаемый объем выборки был бы гораздо меньшим даже с учетом дизайн-эффекта, как было показано в приведенных выше расчетах.

Принятый в бостонском исследовании столь существенный размер (объем) выборки, превышающий величины вышеприведенных расчетов, объясняется тем, что это исследование ставило целью не только оценку урологической заболеваемости⁵ в общей популяции, но также и сравнение этого показателя в возрастном, половом и расово-этническом аспекте [30]. Кроме того, формирование представительной вы-

борки предполагало проведение последующих проспективных популяционных исследований. Данный пример подтверждает положение о том, что усложнение исследования, включение в него дополнительных задач влечет за собой увеличение объема выборки.

Планирование в структуре (дизайне) выборочного исследования редких заболеваний и малочисленных групп населения также приводит к резкому увеличению объема выборки.

Одним из направлений практического использования выборочного метода, которому уделяют все большее внимание, является т.н. методология быстрого обследования (RSM – rapid survey methodology) [12, 31]. Методика, одобренная и рекомендованная Всемирной организацией здравоохранения (ВОЗ), прежде всего для развивающихся стран, первоначально была основана на обследованиях в рамках Расширенной программы иммунизации (EPI), но затем была применена и в исследованиях с другими целями и задачами. Математическое обоснование методики, несмотря на использование в течение ряда лет, было подтверждено лишь недавно. Расчет объема выборки и ее формирование позволяют определить подходящую численность обследования, что обеспечивает необходимую точность и одновременно экономию ресурсов.

При определении объема выборки следует учитывать и т.н. отклик на обследование, показатель которого необходимо стремиться обеспечить как можно более высоким [20].

Таким образом, воспользовавшись накопленным в литературе опытом теории и практики выборочного метода, исследователь может выбрать способ расчета объема выборки населения, ■

³ англ. *stratify* – расслаивать; англ., лат. *stratum* – слой.

⁴ англ. *cluster* – кисть, пучок, гроздь, куст, группа.

⁵ Путем определения распространенности симптомов.

предназначенной для определения показателя общей урологической заболеваемости. При этом необходимо принимать во внимание тип формирования выборки, самой элементарной из которых является простая случайная выборка. Однако следует отметить, что данный тип выборки не всегда применим в популяционных исследованиях

в силу организационных и временных издержек: обычно бывает трудно составить поименный перечень населения определенной территории, как правило, достаточно крупной для такой процедуры, чтобы с этим могла справиться исследовательская группа. Расчет размера выборки, как и ее формирование, требует определенных знаний, навыков,

опыта, а самое главное, точно сформулированных целей и задач исследования.

В ряде случаев, в особенности связанных со сложными масштабными популяционными исследованиями на национальном уровне, формирование выборки целесообразно предоставить специалистам по сэмплингу (выборочному методу) [32]. ■

Ключевые слова: статистика урологических заболеваний, популяционный анализ заболеваемости.

ЛИТЕРАТУРА

- Аполихин О.И., Какорина Е.П., Сивков А.В., Бешлиев Д.А., Солнцева Т.В., Комарова В.А. Состояние урологической заболеваемости в Российской Федерации по данным официальной статистики. // Урология. 2008. № 3. С.3-9.
- Miller D.C., Saigal C.S., Litwin M.S. The demographic burden of urologic diseases in America // Urol. Clin. North Am. 2009. Vol. 36, № 1. P.11-27.
- Anger J.T., Saigal C.S., Wang M., Yano E.M. Urologic disease burden in the United States: veteran users of Department of Veterans Affairs healthcare. // Urology. 2008. Vol. 72, № 1. P.37-41.
- Зиборова И.В., Лопаткин Н.А., Сивков А.В. Социально-экономические аспекты федеральной целевой программы «Урология». // Экономика здравоохранения. 1999. № 4. С. 5-8.
- Гайбуллаев А.А., Кариёв С.С. Состояние и перспективы развития урологической помощи в Республике Узбекистан. Инфекции мочевыводящих путей. Ташкент. 2007. С. 25-35.
- McKinlay J.B., Link C.L. Measuring the urologic iceberg: design and implementation of the Boston Area Community Health (BACH) Survey // Eur. Urol. 2007. Vol. 52, № 2. P. 389-396.
- Случанко Н.С. Изучение здоровья населения. Руководство по социальной гигиене и организации здравоохранения. М. Медицина, 1987. С. 278-313.
- Лисицын Ю.П. Образ жизни и здоровье. Руководство по социальной гигиене и организации здравоохранения. М. Медицина, 1987. С. 131-148.
- Джессен Р. Методы статистических обследований. М., Финансы и статистика. 1985. 478 с.
- Йейтс Ф. Выборочный метод в переписях и обследованиях. М., Статистика. 1965.
- Шварц Г. Выборочный метод. Руководство по применению статистических методов оценивания. М., Статистика. 1978. 213 с.
- Bennett S., Woods T., Liyanage W.M., Smith D.L. A simplified general method for cluster-sample surveys of health in developing countries. // World Health Stat. Q. 1991. Vol. 44, № 3. P.98-106.
- Шехтер Е.Г. О методе определения объема выборочной совокупности при изучении заболеваемости населения. Диспансеризация всего населения (проблемы, пути решения). М. 1986. С. 164-169.
- Шиган Е.И. Методика социально-гигиенических исследований. Руководство по социальной гигиене и организации здравоохранения. М. Медицина. 1987. С. 200-278.
- Бессмертный Б.С. Математическая статистика в клинической, профилактической и экспериментальной медицине. М. Медицина. 1967. 304 с.
- Lwanga S.K., Lemeshow S. Sample size determination in health studies. Geneva: «World Health Organization». 1991.
- Abramson J.H., Gahlinger P.M. Computer Programs for Epidemiologists: PEPI Version 4.0. // Am. J. Epidemiol. 2002. Vol. 155, № 8. P. 776-777.
- Oster R.A. An examination of five statistical software packages for epidemiology. // The American Statistician. 1998. Vol. 52, № 3. P.267-280.
- Обучение медицинской статистике. Двадцать конспектов лекций и семинаров. Женева: «Всемирная Организация Здравоохранения, Медицина». 1989. 216 с.
- Роуз Дж., Блэкберн Г., Гиллум Р.Ф., Принеас Р.Г. Эпидемиологические методы изучения сердечно-сосудистых заболеваний. Женева: «Всемирная Организация Здравоохранения, Медицина». 1984. 224 с.
- Vaughan J.P., Morrow R.H. Manual of epidemiology for district health management. Geneva: «World Health Organization». 1989.
- Чаклин А.В., Осечинский И.В. Основные методические принципы эпидемиологического исследования неинфекционных болезней. Эпидемиология неинфекционных заболеваний. М. Медицина. 1990. С. 190-215.
- Сковердяк Л.А. Распространенность заболеваний почек и мочевыводящих путей и объем лечебно-профилактической помощи: Автореф. дисс. канд. мед. наук. М. 1975.
- Савченко Н.Е., Строцкий А.В., Жлоба П.П. Нехирургические методы лечения доброкачественной гиперплазии простаты: (Пособие для практических врачей). Минск. 1998. 63 с.
- Лопаткин Н.А., Зиборова И.В., Сивков А.В., Аполихин О.И. Экономические вопросы лечения доброкачественной гиперплазии предстательной железы. // Урология и нефрология. 1999. № 1. С.22-26.
- Resnick M.I., Persky L. Summary of the National Institutes of Arthritis, Diabetes, Digestive and Kidney Diseases conference on urolithiasis: state of the art and future research needs. // J. Urol. 1995. Vol. 153, № 1. P.4-9.
- Кожобеков Б.С. Комплексное изучение эпидемиологических и этиологических аспектов, ранней диагностики, оценки лечебных пособий и метафилактики мочекаменной болезни: Автореф. дисс. докт. мед. наук. Алматы. 1998.
- Plata R., Silva C., Yahuita J., Perez L. et al. The first clinical and epidemiological programme on renal disease in Bolivia: a model for prevention and early diagnosis of renal diseases in the developing countries. // Nephrol. Dial. Transplant. 1998. Vol.13, № 12. P. 3034-3036.
- Актуальные проблемы профилактики неинфекционных заболеваний. // Материалы научной конференции. «Профилактика заболеваний и укрепление здоровья». 2004. С. 3.
- Kupelian V., Wei J.T., O'Leary M.P. et al. Prevalence of lower urinary tract symptoms and effect on quality of life in a racially and ethnically diverse random sample: the Boston Area Community Health (BACH) Survey. // Arch. Intern. Med. 2006. Vol. 166, № 21. P. 2381-2387.
- Frerichs R.R., Tar K.T. Computer-assisted rapid surveys in developing countries. // Public Health Rep. 1989. Vol. 104, № 1. P. 14-23.
- Медико-демографическое исследование Узбекистана 1996. Ташкент Калвертон: «Институт акушерства и гинекологии Министерства здравоохранения РУз; Макро Интернэшнл». 1997.