

КАК ЧИТАТЬ МЕДИЦИНСКИЕ СТАТЬИ: ЧАСТЬ 2. ИССЛЕДОВАНИЯ, ПОСВЯЩЕННЫЕ МЕТОДАМ ДИАГНОСТИКИ

В.В. Власов

Использование диагностических методов в медицине весьма разнообразно. Теория их применения и практические рекомендации по оценке эффективности детально описаны [1]. Остановимся лишь на том, что совершенно необходимо для понимания и оценки публикаций о методах диагностики.

Диагностический тест (прием, метод) может быть использован в разных целях.

А. Если исследование проводят людям, считающим себя здоровыми, с целью выявить тех, кто болен, хотя не подозревает об этом, такое исследование называется *скринингом* (от англ. screen — просеивать). Пример — массовые флюорографические исследования.

Б. Если исследование проводят всем людям, оказавшимся в больнице или поликлинике в связи с любым заболеванием, то это называется *поиском сопутствующих заболеваний* (case finding). Примеры — общий анализ мочи, клинический анализ крови, обычно назначаемые в подобных случаях.

В. Если исследование проводят людям, обратившимся с жалобами, для выяснения причины их страданий, уточнения диагноза, то это — собственно *диагностическое исследование*, самый важный случай. Примеры — электрокардиография при жалобах на перебои в работе сердца, колоноскопия при жалобах на следы крови на каловых массах.

Г. Если исследование проводят людям, получающим лечение, чтобы оценить его успешность, то это — тест на эффективность лечения (*мониторинг*). Пример — термометрия у больного с пневмонией, получающего антибиотики.

Журнальную статью о методе диагностики, которая привлекла Ваше внимание и прошла первые этапы критической оценки, следует рассмотреть детально с применением критерии, обобщенных в табл. 1.

Сопоставлен ли предлагаемый авторами метод с "золотым стандартом" (референсным методом), то есть с самым надежным из соответствующих методов? Разрабатывая новый метод, авторы обычно имеют в виду, что он обеспечивает определенные преимущества перед теми, которые уже применяются. Может быть, диагностика становится прижизненной, более ранней, более щадящей, более дешевой, в то время как метод "золотого стандарта" точен, но его использование ограничено, (например, это интраоперационный или патологоанатомический метод,

Таблица 1.

Вопросы для оценки исследования по изучению диагностического теста (по [2], с изменениями)

1. Есть ли сопоставление с "золотым стандартом"?
2. Проведено ли это сопоставление слепым методом?
3. Четко ли описан способ формирования изученной группы пациентов?
4. Имели ли обследованные пациенты достаточно широкий спектр заболевания (легкие и тяжелые случаи, леченющиеся и нелечимые, с поражениями разной формы, с сопутствующими заболеваниями)?
5. Оценена ли воспроизводимость результатов теста?
6. Оценена ли надежность интерпретации результатов слепым методом?
7. Рационально ли определена "норма" — отсутствие болезни?
8. Если тест предлагается как дополнение к пакету других тестов, то показано ли, что применение этого теста улучшает общий результат диагностики?
9. Обоснована ли "точка разделения" здоровых и больных?
10. Достаточно ли детально описана тактика применения теста, чтобы Вы могли его правильно использовать?
11. Оценена ли общая полезность теста?

или же длительное наблюдение за развитием поздних проявлений). Обязательная задача исследования диагностического метода — оценить, насколько хорошо диагноз с помощью нового метода соответствует "золотому стандарту". Если авторы не провели такого сравнения — отбрасывайте статью.

Если сравнение с "золотым стандартом" присутствует, то оцените примененные тесты. Удовлетворителен ли референсный метод? Действительно ли авторы использовали наиболее надежный метод диагностики, рассматриваемый медицинской как стандартный? Иногда взятый авторами референсный метод сам по себе недовлетворителен. Тогда не следует тратить время, чтобы разобраться, соответствует ли ему новый тест.

Убеждаются ли Вас резоны авторов в пользу предлагаемого ими метода диагностики? Возможно, он проще, более приемлем для пациента, дешевле? Если он столь же труден или малодоступен, как референсный, тогда его применение не облегчит и не улучшит Вашу практику.

Адрес для корреспонденции:

Проф. В.В. Власов, Саратовский медицинский университет,
410601 Саратов, а/я 1528

Традиционный метод оценки результатов теста — построение так называемого латинского квадрата (четырехпольной таблицы, рис. 1). В статье должна быть приведена такая таблица или сведения, достаточные для ее создания. В противном случае отбрасывайте статью — она не позволит Вам узнать, полезен ли рассматриваемый диагностический тест.

Рассмотрим в качестве примера диагностический эксперимент [3], когда двум группам специалистов — иридодиагностам и офтальмологам было предложено провести диагностику заболеваний почек по радужной оболочке глаза. Нарушение выделительной функции было подтверждено клинически и биохимически (референсный метод). Испытание проводилось слепым методом по цветным слайдам, что принято в иридиологии. Данные, представленные на рис. 1, позволяют вычислить так называемые операционные характеристики диагностического теста (табл. 2): чувствительность, специфичность и диагностическую ценность положительного и отрицательного результатов теста. Табл. 2 показывает, что по радужной оболочке глаза выявить заболевания по-

Таблица 2.

Операционные характеристики метода иридодиагностики болезней почек в случае его применения иридодиагностами (I) и офтальмологами (II)

	I	II
Чувствительность D/(B+D)	61%	44%
Специфичность A/(A+C)	38%	62%
Прогностическая ценность положительного результата (ПЦП)=D/(C+D)	33%	37%
Прогностическая ценность отрицательного результата (ПЦО)=A/(A+B)	65%	69%

чек не способны ни иридодиагносты, ни оphthalmологи (последние на это и не претендуют; они лишь справедливо были избраны в качестве контрольной группы специалистов). В самом деле, прогностическая ценность положительного и отрицательного результата тестируемания (то есть заключения о наличии или об отсутствии болезни соответственно) как иридодиагностами, так и офтальмологами слишком низка.

Интересно, что, хотя обе группы специалистов располагали одинаковой информацией, иридодиагносты констатировали болезнь гораздо чаще, чем офтальмологи (чувствительность теста 61% и 44% соотв.). Это типично для практики "целителей" и "неортодоксальных" диагностов — они находят болезни почти у каждого, причем обычно не одну, а больше. Поскольку отсутствие "ранней стадии" болезни доказать практически невозможно, пациент становится привязанным к такому знаанию и, естественно, лечение "помогает" — ведь болезни-то на самом деле нет! Открывается простор для внушения и развития зависимости пациента от знахаря. Серьезные врачи-исследователи редко тратят время на изучение подобных методов. Именно поэтому так мало добрых данных о "нетрадиционных" методах. Исследования же, проводимые адептами этих методов, обычно не выдерживают критики.

Какие факторы нужно учитывать при рассмотрении результатов теста, рассчитываемых на основе латинского квадрата (табл. 2)? Чувствительность и специфичность называют стабильными характеристиками диагностического теста. Это означает, что они не зависят от распространенности болезни в исследуемой группе. Если чувствительность равна 0,75 (75%), то независимо от того, сколько больных этой болезнью находится в исследуемой Вами группе, 3/4 из этих больных будут выявлены с помощью теста. Если специфичность 0,9, то среди лиц, не имеющих данного заболевания, положительный результат (ложный) будет получен лишь у 10%.

При выполнении диагностического теста врача в основном интересует, насколько высока вероятность болезни у лиц с положительным результатом и насколько она низка у лиц с отрицательным результатом. Этот показатель называется прогностической

I			
	Здоровые, креатинин крови <1,2 мг%	Больные, креатинин крови >1,5 мг%	Всего
Почки здоровы (отрицательный результат)	36 (ИО) A	19 (ЛО) B	A+B=55
Есть болезнь (положительный результат)	C (ЛП) 59	D (ИП) 29	C+D=88
Всего	A+C=95	B+D=48	143

II

	Здоровые, креатинин крови <1,2 мг%	Больные, креатинин крови >1,5 мг%	Всего
Почки здоровы (отрицательный результат)	(ИО) A	27 (ЛО) B	A+B=86
Есть болезнь (положительный результат)	C (ЛП) 36	D (ИП) 21	C+D=57
Всего	A+C=95	B+D=48	143

Рисунок 1.

Усредненные данные эксперимента по диагностике заболеваний почек по радужке глаза опытными иридодиагностами (I) и офтальмологами (II) (по [6] с изменениями).

Варианты совпадения и несовпадения референсных данных и результатов диагностического теста: ИО — истинно-отрицательные случаи, когда истинное отсутствие заболевания совпало с отрицательным результатом теста (здоров); ЛО — ложноотрицательные случаи, когда у больных людей получен отрицательный результат (здоров); ЛП — ложно-положительные случаи, когда истинное отсутствие заболевания совпадает с положительным результатом теста (болен); ИП — истинноположительные случаи, когда истинная болезнь совпадает с положительным результатом (болен).

ценностью, положительных (ПЦП) и отрицательных (ПЦО) результатов. Почему этими показателями нельзя ограничиваться? Потому, что их величина зависит от распространенности заболевания в изучаемой группе. При чувствительности и специфичности 90% (это высокие показатели) тестирование в условиях низкой распространенности (когда в исследуемой группе только 10% больных) даст ПЦП=50% ПЦО=99%, тогда как при высокой распространенности (в группе 80% больных) ПЦП=97% и ПЦО=69%. Следовательно, даже если автор статьи делает упор на высокой надежности окончательного диагноза (прогностической ценности положительного результата), надо смотреть, каковы чувствительность и специфичность. Возможно, привлекательный результат — просто следствие того, что в исследуемой группе распространенность заболевания очень высока, например, благодаря предварительному отбору пациентов.

Тест, привлекательный в условиях клинической диагностики, может быть непригоден для скрининга. Яркий пример — электрокардиография. Незаменимая при обследовании больных с возможными заболеваниями сердца, она практически бесполезна в профилактических обследованиях населения. Нагрузочное ЭКГ-тестирование — весьма эффективный метод для обследования мужчин среднего возраста, предъявляющих жалобы на боли в груди. Но этот метод не информативен в условиях низкой распространенности ишемической болезни сердца (ИБС) — при профилактических обследованиях молодых мужчин и женщин. У пожилого мужчины, предъявляющего жалобы, соответствующие типичной стенокардии, нагрузочное ЭКГ-тестирование тоже практически бесполезно — оно не может ни заметно увеличить вероятность ИБС при положительном его результате, ни существенно снизить вероятность болезни при отрицательном. Это означает, что нагрузочное ЭКГ-тестирование (впрочем, как и любой другой диагностический тест) максимально полезно в условиях неопределенности диагноза.

Применен ли для сопоставления с "золотым стандартом" слепой метод? При правильной организации исследования оценка результатов изучаемого теста должна быть произведена "вслепую" — так, чтобы специалист не знал результатов других анализов у этого пациента, прежде всего — результатов референсного теста. Его результаты тоже должны оцениваться вслепую, независимо от результатов изучаемого теста. В противном случае происходят ошибки типа "смещения к предполагаемому диагнозу". Обычно они возникают непреднамеренно; зная о предположительном диагнозе, специалист склонен к предвзятости при оценке результатов (данных рентгенографии, ЭКГ и т.д.).

Четко ли описан способ формирования исследуемой группы пациентов? Должно быть указано, по какому принципу пациенты подбирались, каковы критерии включения и исключения, какова исходная совокупность — посетители обычной поликлиники или больные специализированного стационара. От этого зависит не только распространенность болезни в исследуемой группе, но и ее тяжесть, соотношение разных форм болезни и многое другое.

Если группа "контрольных" пациентов (не имеющих данной болезни) подбиралась отдельно, то столь же ясно должно быть сказано, как это делалось. Иначе может случиться, что отличия "больных" от "здоровых" обусловлены тем, что " здоровые" — это молодые активные студенты, а "больные" — пациенты пожилого возраста в стационаре.

Достаточно ли широк спектр тяжести патологии у обследованных? Адекватно ли представлены в контрольной группе сопутствующие и сходные заболевания, а в группе больных по референсному критерию — разные формы основного заболевания? Возможно, полученные авторами характеристики изучаемого теста относятся лишь к одной, тяжелой форме заболевания или к его поздней стадии, или только к случаям, когда отсутствуют другие сходные заболевания. Допустим, тест позволяет отличить здоровых от больных раком толстой кишки, но не помогает отличить больных раком от больных колитом. Поскольку в обычной практике встречаются все формы болезни и любые сопутствующие заболевания, такой тест окажется практически бесполезным.

Оценена ли воспроизводимость результатов диагностического теста? Хороший тест должен быть воспроизводимым и точным. Следует помнить, что воспроизводимость результатов инструментальных и, тем более, неинструментальных клинических исследований далеко не так высока, как обычно представляется. Особенно низка воспроизводимость результатов ультразвуковых, рентгенологических, радиоизотопных, электрокардиографических и других исследований, связанных с оценкой изображения. Большинство врачей переоценивают надежность эндоскопической диагностики, считая ее почти абсолютной.

Надежность интерпретации результатов теста должна быть оценена слепым методом. Это означает, что результат оценивается повторно тем же специалистом (сходимость) или разными (воспроизводимость) в условиях, когда эти специалисты не имеют никаких других данных о пациенте.

Рационально ли определена "норма" — отсутствие болезни? При испытании диагностических тестов обычно сравнивают больных не со " здоровыми", а с теми, у кого, согласно референсному диагностическому методу, нет изучаемой болезни. В некоторых случаях исследователь сопоставляет результат теста у "больных" с результатом в "норме". Тогда должно быть ясно сформулировано, что имеется в виду под "нормой".

Описание и уточнение понятия "нормы" в клинической практике посвящена специальная область исследований. Концепция "нормы" используется очень широко и она весьма расплывчата. Почти любое толкование "нормы" имеет свои недостатки.

В лабораторном деле отказались от термина "норма" в пользу более конкретных "референсных величин". Если подобрать группу людей, отвечающих некоторым стандартным требованиям клинического здоровья, то такую группу можно назвать референсной. Применение теста к людям, составляющим такую группу, в некоторых стандартных условиях и при стандартной методике дает совокупность референсных величин. Эти референсные величины в стати-

стическом смысле распределены некоторым образом, их распределение можно описать с помощью известных математических способов. Например, многие фактические распределения удовлетворительно описываются с помощью средней величины и среднего квадратического отклонения, то есть они соответствуют закономерностям "нормального" или Гауссского распределения. В случае, если фактическое распределение не соответствует какому-либо теоретическому или если исследователь принципиально не желает искать способ описания, он может найти границы, включающие, например, 95% средних величин (или 90%, 99%), методом персентилей [1]. Такой интервал, содержащий большинство референсных величин, называют референсным интервалом. Референтный интервал — это не норма "вообще", а характеристика определенной референсной группы и условий исследования. Выходящее за его пределы не считается "ненормальным", оно лишь нехарактерно для референсной группы.

Важно подчеркнуть, что при любом подходе к определению границ "нормы" применение этого понятия на практике приводит к тому, что почти у всех здоровых людей обнаруживаются какие-нибудь "отклонения" [4]. Чем больше разнообразных тестов, тем больше отклонений. В итоге — лишие затраты значительных средств на дополнительное обследование людей, у которых выявлены "отклонения". В особенности остро встает эта проблема при скрининге, то есть в условиях низкой распространенности заболевания. Если границы нормы определены произвольно, то и результат произведен. Станем, например, считать повышенной массы тела выше средней; тогда половина населения окажется с "повышенной" массой тела.

Когда границы "нормы" устанавливают исходя из распределения величин у "здоровых" людей, то игнорируют различия заболеваний между собою. В действительности при исследовании, например, числа лейкоцитов в крови при одних болезнях получаются величины более близкие к "норме", а при других — более отличающиеся. Поэтому для разных болезней следует устанавливать свой диагностический порог (точку разделения — см. ниже). Он может находиться как внутри "нормы", так и за ее пределами.

С позиций выбора критерии эффективного вмешательства необходимо выделить величину, начиная с которой лечение оказывается оправданным. Простой пример — гипертония. Величина артериального давления 140/90 мм рт.ст. — распространенный критерий диагностики. Однако этот уровень не рассматривается как критерий для решения о медикаментозном лечении.

С величиной отдельных показателей связан риск развития будущих заболеваний. Поэтому соблазнительно связывать определение "нормы" с величиной риска болезни и смерти. А тогда неизбежно возникают противоречия, поскольку низкий риск одного заболевания может сочетаться с повышенным риском другого. Здоровье предстает лишь как некоторый компромисс.

Границы медицины исторически изменичивы. С течением времени от шаманства, изгнания бесов и

сожжения ведьм общество идет к более четкому ограничению сферы приложения медицины. Такие явления, как злоупотребление психиатрией в политических целях, можно рассматривать в свете экстраполяции медицинских представлений о норме на поведение людей, его приемлемость в рамках определенной культурной, социальной и политической системы.

Злоупотребление понятием нормы имеет и еще один аспект. Иногда нормой считают активное, десспособное состояние, увязывающее ее с профессиональными возможностями человека. В действительности это не относится к прерогативе медицины.

Обычно в клинической практике тест применяется в совокупности с другими. Тем не менее, полезность изучаемого теста должна быть продемонстрирована при его самостоятельном применении. В некоторых исследованиях при очевидной неэффективности диагностики с помощью изучаемого теста предлагается использовать его в качестве "дополнения к другим методам исследования". Это ложная рекомендация. Тест, неэффективный сам по себе, ничего не добавит к другим.

При использовании теста в совокупности с другими должно быть определено, в каком случае, применительно к каким больным, при каких результатах других тестов применяется этот тест. Должна быть показана его эффективность в выделенной другими средствами группе больных. В заключение должно быть оценено, насколько дополнительное применение этого теста улучшило диагностику в целом. Только такая оценка эффективности пакета тестов доказательна.

Обоснован ли выбор "точки разделения" здоровых и больных? Точкой разделения называют величину физиологического показателя, которая служит границей, разделяющей здоровых и больных (годных и негодных, нуждающихся в дальнейшем обследовании и не нуждающихся). Пример точки разделения — концентрация глюкозы в венозной крови 5,83 ммоль/л (ферментный метод). В случае обнаружения натощак более высоких концентраций считается необходимым проводить дополнительные исследования (например, нагрузочную сахарную пробу), чтобы выявить или исключить диабет [3].

В действительности при концентрации глюкозы $>5,83$ ммоль/л вероятность диабета не возрастает скачкообразно [5]. На рис. 2 видно, что у здоровых людей и у больных диабетом концентрация глюкозы значительно варьирует. При описании этих распределений говорят, что они широко перекрываются. Выбрав точку разделения (x_2), можно оценить чувствительность, специфичность и прогностическую ценность положительного и отрицательного результатов. Точку разделения можно сместить в сторону более высоких концентраций (x_3). Тогда чувствительность снизится, но зато возрастет специфичность и прогностическая ценность положительного результата. Если сдвинуть точку разделения влево (x_1), то повысится чувствительность и прогностическая ценность отрицательного результата, но снизится специфичность и прогностическая ценность положительного результата.

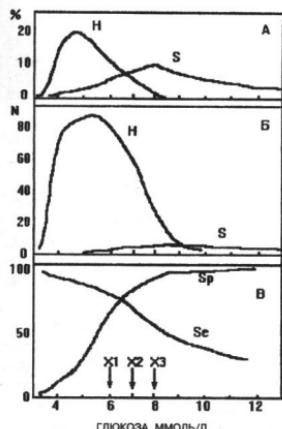


Рисунок 2.

Распределение концентрации глюкозы у здоровых людей (H) и больных диабетом (S) при исследовании венозной крови натощак.

А — распределение обследованных в %; Б — распределение в абсолютных цифрах; В — изменения чувствительности (Se) и специфичности (Sp) при сдвиге точки разделения (x_1 , x_2 , x_3) (по [5], с изменениями).

Чтобы описать, как изменяется результат исследования в зависимости от выбора точки разделения (то есть в зависимости от того, насколько придирчиво врач оценивает рентгенограмму или другой результат), используют так называемый ROC-анализ (Receiver Operating Characteristic analysis). ROC-кривая или характеристическая кривая — это кривая зависимости чувствительности от вероятности ложноположительных результатов, то есть величины 1-специфичность (рис. 3). Чем ниже опускается точка разделения по шкале концентрации глюкозы, тем полнее выявляются больные. Одновременно все большее число здоровых людей оцениваются как больные.

Информативность диагностического теста определяется тем, насколько высоко лежит его характеристическая кривая. Выбор точки разделения не меняет положения кривой. Происходит лишь перемещение по кривой. Наиболее простой и понятный способ описания этой кривой — измерение площади под кривой. Чем ближе площадь под кривой к площади квадрата, тем эффективнее диагностический тест (кривая 1 на рис. 3). Чем ближе кривая к диагонали, тем ниже эффективность диагностического метода (кривая 2 на рис. 3). Диагональ соответствует полной неразличимости здоровых и больных изучаемым методом. Точки, лежащие ниже диагонали, соответствуют «обратной» диагностике, то есть ситуации, когда здоровые чаще называются больными, чем истинные больные. На рис. 4 приведены результаты попытки диагностировать заболевания почек по радужной оболочке глаза [6]. Хорошо видно,

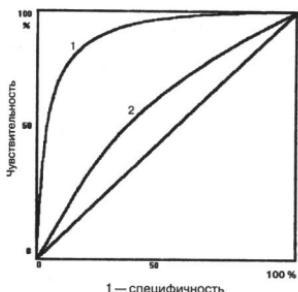


Рисунок 3.

Характеристические кривые для высоконформативного (1) и низконформативного (2) методов диагностики

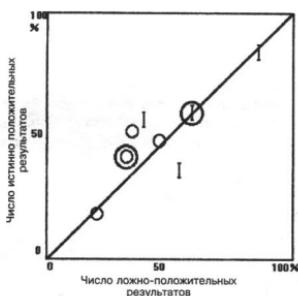


Рисунок 4.

Эффективность диагностики заболеваний почек по радужной оболочке глаза приодиагностами (I) и офтальмологами (O). Кружком обведены средние величины (по данным [6]).

что все экспериментальные точки сосредоточились около диагонали, значит метод неэффективен.

Построение характеристической кривой — наиболее современный и полный метод описания эффективности диагностического теста. Когда имеется характеристическая кривая, можно обсуждать вопрос о том, какую точку разделения выбрать. Разные подходы к выбору описаны детально [1, 7]. Важно подчеркнуть, что простые способы не дают хороших результатов. Например, ошибочно стремление к минимуму числа ложных результатов. При таком подходе уравниваются ложноположительные и ложноотрицательные результаты. Между тем, цена диагностических ошибок почти никогда не бывает одинаковой. Ложноположительный результат может угрожать, например, лишь дополнительным визитом к врачу, а ложноотрицательный — невыявлением опасной, но излечимой болезни. Последствия ошибок и правильных решений — свои для каждой конкретной ситуации. Именно поэтому выбор точ-

ки разделения через взвешивание последствий — очень ответственное дело. К сожалению, сегодня не существует универсальных рецептов для этого. Изучайте внимательно, насколько обоснован выбор точки разделения в каждом случае.

Детали применения теста должны быть подробно описаны. Как надо готовить пациента? Как надо обрабатывать кожу? Как хранить и транспортировать пробы? Неясность в подобных "простых" деталях может сделать выполнение теста невозможным.

Достаточно ли детально объяснена тактика применения теста, чтобы Вы могли его использовать? Ясен ли подбор пациентов? Всем ли можно и нужно проводить такое исследование? Или же для каких-то групп пациентов тест неэффективен, опасен? Каковы опасности и как их избежать?

Оценена ли общая полезность применения теста? Критерий эффективности диагностики — возможность выяснения истинного диагноза. Критерий полезности теста — улучшение исхода лечения. Не каждая статья о диагностическом методе содержит информацию о его полезности, но от этого зависит Ваше решение, применять ли метод.

Каковы преимущества пациентов, которым установлен диагноз? Существует ли при данном заболевании эффективное лечение? Какова судьба пациентов с ложноположительным диагнозом? Каковы последствия для пациентов с ложноотрицательным диагнозом? Существующее в медицине принципи-

альное стремление к точности диагноза нельзя признать всегда оправданным, поскольку во многих случаях специфическое лечение отсутствует и проводится лечение неспецифическое. Поэтому иногда уточнение диагноза бесполезно, ибо откладывает лечение и увеличивает страдания и расходы.

Сложившаяся практика — еще не доказательство полезности действий. В особенности это относится к так называемой профилактической деятельности врача. Например, в течение многих десятилетий врачи оценивают "физическое развитие" взрослых людей, хотя это не имеет ясного смысла, методика расплывчатая и оценка не несет никаких результатов, если не считать таковыми заполнение отчетных таблиц [8].

Несмотря на то, что в ведущих журналах рецензенты применяют к статьям описанные выше критерии, в большинстве публикемых статей Вы не найдете исчерпывающих ответов на все вопросы, уже хотя бы потому, что достаточные сведения о диагностическом методе накапливаются в ходе множества исследований. Кроме того, некоторые детали могут быть сообщены в отдельных публикациях. Например, результаты оценки воспроизводимости метода иногда публикуются отдельно от результатов его оценки в сопоставлении с референсным методом. Однако в любом случае при анализе статьи исходите из соответствующих критерии. Только так Вы будете узнавать и вносить в свою практику лучшие, проверенные методы диагностики.

Литература

1. Власов В.В. Эффективность диагностических исследований. М 1988:245.
2. How to Read Clinical Journals: 2. To Learn About a Diagnostic Test. Can Med Ass J 1981;124:703—710.
3. Тиш Н.В. Клиническая оценка лабораторных тестов. М 1986:486.
4. Власов В.В., Никольский В.А. Вероятность отклонения от референтных интервалов при скрининге по множеству параметров. Лаб дело 1984;1:46—49.
5. Remein Q.R., Wilkerson H.L. The efficiency of screening tests for diabetes. J Chron Dis 1961;13:6—21.
6. Simon A., Worthen D.M., Mitas J.A. An evaluation of iridology. JAMA 1979;242:13:1385—1389.
7. Ластед Л. Введение в проблему принятия решений в медицине. Пер. с англ. М 1971:282.
8. Власов В.В. Следует ли оценивать физическое развитие военнослужащих? Воен мед журн 1993;3:8—11.