International Journal of Multidisciplinary Research and Development
www.allsubjectjournal.com
ISSN Online: 2349-4182 ISSN Print: 2349-5979
Received: 25-08-2022, Accepted: 10-09-2022, Published: 26-09-2022
Volume 9, Issue 9, 2022, Page No. 58-63

Assessment of the exposure renal failure patients using diagnostic tests accuracy

Ali Ibrahim Ali Elkhidir, Altaiyb Omer Ahmed Mohmmed
Department of Statistics, Sudan University of Science and Technology, Khartoum, Sudan

Abstract
Clinically, it is always important to know how good the test is at predicting the true disease status given the findings from new proposed test, the essential aim of this paper is to provide the basic conceptual framework and interpretation of components, analysis, and comparison of diagnostic accuracy tests. the observed exposure–outcome using popular measures such as sensitivity, specificity, positive predictive value, negative predictive value, diagnostic odds ratio, risk ratio, positive likelihood ratio, negative likelihood ratio, and the accuracy of a diagnostic test. Data include only those RF patients who are on dialysis or have received a kidney transplant. Questionnaire was used to get a data, 94 patients whom treated, and follow-up study was used to offer the data. most diagnostic test calculated of Chronic Kidney Disease risk factors was significantly associated among the people studied regarded to gender and age, for survive a person was the life is more likely in the younger rather older, a two times dialysis per week for both male and female, also for those suspected of (HTN) is more likely than (DM), for duration of disease, a person more than 5 years was more likely to survive than whom less than 5 years.

Keywords: CKD, confounding diabetes, likelihood ratio, sensitivity

Introduction

Clinically, it is always important to know how good the test is at predicting the true disease status given the findings from new proposed test, many diagnostic tests and procedures are applied for the diagnosis of a suspected Renal Failure (RF) Different measures may be used to describe how often disease or another health event occurs in a population. The performance of these tests can be investigated in diagnostic accuracy studies. A fundamental aspect of the evaluation of diagnostic tests is test accuracy, that is, the ability of a test to differentiate between those who have and those who do not have the condition or disease of interest (Sitch *et al* 2021) [31]. The key element in accuracy studies is the application of a gold standard test determining the true disease state of each patient, discriminative measures are mostly used by health-policy decision makers; predictive measures are most useful for predicting the probability of a disease in an individual. To use a diagnostic test effectively and consistently in their practice (Thomas *et al* 2018) [27], clinicians need to know how well the test distinguishes between those patients who have the suspected acute or chronic disease and those patients who do not.

What is also important is the fact that measures of a test performance are not fixed indicators of a test quality. On the contrary, measures of diagnostic accuracy are very sensitive to the characteristics of the population in which the test accuracy is being evaluated. Some measures largely depend on the disease prevalence, whereas others are highly sensitive to the spectrum of the disease in the studied population. It is therefore of utmost importance to understand the meaning of different measures of diagnostic accuracy and to know how to interpret them and under what conditions they may be used (Simundic 2012) [23].

1. Measures of Test Accuracy

Diagnostic accuracy measures tell us about the ability of a test to discriminate between and/or predict disease and health. This discriminative and predictive potential can be quantified by measures of diagnostic accuracy such as sensitivity and specificity, predictive values, likelihood ratios, area under the receiver operating characteristic curve, overall accuracy and diagnostic odds ratio (Eusebi 2013 p267) [21].

Test accuracy is determined by cross classifying the results (positive and negative) of an index test against those of the reference standard (Takwoingi *et al* 2013) [6]. Traditionally, the 2 ×2 table is used to explain the different measures associated with accuracy for dichotomous measures. Valid estimates of comparative accuracy can be obtained from comparative diagnostic test accuracy (DTA) studies: a comparison of accuracy of two or more tests (none of which is the reference standard) within a single study (Leeflang, & Reitsma 2018) [7].

The interpretation of a diagnostic test result depends on both the ability of the test to distinguish diseased from non diseased subjects and the particular characteristics of the patient and setting in which the test is being used (Zweig & Campbell 2018).

2. Measures of Diagnostic Accuracy

The discriminative (Eusebi 2013 p268) [21] ability of a test can be quantified by several

- Sensitivity and Specificity;
- Positive and Negative Predictive Values (PPV, NPV);
- Positive and Negative Likelihood Ratios (LR+, LR–);
- The Area Under curve (AUC).
- The Receiver Operating Characteristic (ROC);
- The Diagnostic Odds Ratio (DOR);
- The Overall Diagnostic Accuracy.
- These measures are related to two main categories of issues:
- Classification of people between those who are and those who are not diseased (discrimination);
- Estimation of the posttest probability of a disease (prediction).

Table 1: illustrates how diagnostic test can be computed if the result in each patient has been classified into one of the four categories True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN)

	Test result	**True Diagnosis**	Total
	Present (D+)	**Absent (D-)**	
Positive (T+)	TP	FP	TP+FP
Negative (T-)	FN	TN	FN+TN
Total	TP +FN	FP+ TN	N

Where
TP = number of true positive results, FP = number of false positive results,
FN = number of false negative results, TN = number of true negative results,
TP +FN = number of positive results of the reference test, FP+ TN = number of negative results of the reference test, TP+ FP = number of positive results of the index test, FN+ TN = number of negative results of the index test, N = overall sample size.
Here a positive result of the index test is called true positive, if the target condition is actually present in the patient– as indicated by the reference test – and false positive, if the target condition is actually not present. A negative result of the index test is called true negative, if the target condition is actually not present in the patient, and false negative, if the target condition is actually present (Vach *et al* 2017) [13].

2.1 Sensitivity
The sensitivity (Lalkhen & McCluskey 2008) [22] of a clinical test refers to the ability of the test to correctly identify those patients with the disease.
Sensitivity [Altman &, Bland 1994] [10] is generally expressed in percentage and defines the proportion of TP subjects with the disease in a total group of subjects with the disease:

$$Sens = \frac{TP}{TP + FN} \quad (1)$$

Sensitivity estimates the probability of getting a positive test result in subjects with the disease. Hence, it relates to the ability of a test to recognize the ill.

2.2 Specificity
The specificity of a clinical test refers to the ability of the test to correctly identify those patients without the disease.
In the other hand, Specificity [Naaktgeboren *et al* 2016], is defined as the proportion of subjects without the disease with a negative test result in a total group of subjects without the disease:

$$Spec = \frac{TN}{TN + FP} \quad (2)$$

In other words, specificity estimates the probability of getting a negative test result in a healthy subject. Therefore, it relates to the ability of a diagnostic procedure to recognize the healthy (Eusebi 2013) [21].

2.3 Positive Predictive Value
The PPV is of particular importance and is an often- used measure to quantify the degree of information bias, and hence to quantify the degree of misclassification [Newcomer *et al* 2019] [8].
The PPV is the probability that a patient has the disease given that the test results are positive; therefore, a PPV represents a proportion of patients with a positive test result in a total group of subjects with a positive result: [Thompson & Bruel 2011] [25]

$$PPV = \frac{TP}{TP + FP} \quad (3)$$

2.4 Negative Predictive Value
The NPV is the probability that a patient does not have the disease given that the test results are indeed negative. NPV describe the probability of not having a disease for a subject with a negative test result. An NPV is defined as a proportion of subjects without the disease and with a negative test result in a total group of subjects with a negative test result: [Bolin & Lam 2013] [28].

$$NPV = \frac{TN}{FP + FN} \quad (4)$$

2.5 Likelihood Ratio
Likelihood ratios can be used to help adapt the results of a study to your patients. This is defined as how much more likely is it that a patient who tests positive has the disease compared with one who tests negative.
LR is useful measures of diagnostic accuracy, but they are often disregarded, even if they have several particularly powerful properties that make them very useful from a clinical perspective. They are defined as the ratio of the probability of an expected test result in subjects with the disease to the probability in the subjects without the disease [Deeks & Altman 2004] [11].
We distinguish (Vach *et al* 2017 p106) [13] between the positive likelihood ratio (LR+) and the negative likelihood ratio (LR-). Both measures are based on sensitivity and specificity and are independent of the prevalence of the disease. Based on sensitivity and specificity the LRs are defined as:

$$LR+ = \frac{TP(FP + TN)}{FP(TP + FN)} \quad (5)$$

And

$$LR- = \frac{FN(FP + TN)}{TN(TP + FN)} \quad (6)$$

The magnitude of the LR informs about the certainty of a positive diagnosis. As a general guideline, a value of LR =1 indicates that the test result is equally likely in patients with and without the disease, values of LR >1 indicate that the test result is more likely in patients with the disease and values of LR <1 indicate that the test result is more likely in patients without the disease (Zhou *et al* 2002) [9].
LR+ denotes the ratio of the probability of a positive test result among diseased and the probability of a positive test result among non-diseased. LR- describes the ratio of the probability of a negative test result among diseased and non-diseased, respectively. Both LRs range from 0 to plus infinity. An LR+ or LR- of one indicates that the particular test result is equally likely in the diseased and non-diseased population. An LR+; LR- greater than one indicates that the test result is more likely in the diseased than in the non-

diseased. For a reasonable test LR+ is expected to be much greater than one, and LR- is expected to be much less than one (Xue & Goswami 2019) [26].

2.6 Youden Index

The Youden index (Youden, 1950) [12] is another univariate measure of accuracy, defined as:

$$J = sens + spec - 1 \quad (7)$$

The Youden index ranges from -1 to 1. Values close to zero occur if a diagnostic test gives the same proportion of positive results for diseased and non-diseased, i.e the test is useless.

Values close to 1 indicate that the test performs well and is able to distinguish between the two disease states. As the index gives equal weight to false positive and false negative values, diagnostic tests with the same value show the same proportion of total misclassified results (Sergeant *et al* 2021) [30].

2.7 Diagnostic Odds Ratio

The diagnostic odds ratio (DOR) of a test is the ratio of the odds of a positive test for a subject who has the disease relative to the odds of a positive test for a subject who is disease-free. What is an 'odds'? Any number (e.g., a probability or an observed relative frequency) (p between 0 and 1) can be transformed into its 'odds', which is p/(1 - p) and lies between zero and plus infinity (Vach *et al* 2017 p108) [13].

$$DOR = \frac{TP \times TN}{FP \times FN} \quad (8)$$

2.8 Risk Ratio

A risk ratio (RR), also called relative risk, compares the risk of a health event (disease, injury, risk factor, or death) among one group with the risk among another group. It does so by dividing the risk (incidence proportion, attack rate) in group 1 by the risk (incidence proportion, attack rate) in group 2. Often, the group of primary interest is labeled the exposed group, and the comparison group is labeled the unexposed group (Arias *et al* 2003) [1].

$$RR = \frac{TP(FN + TN)}{FN(TP + FP)} \quad (9)$$

3. Statistical Analysis

The purpose of a diagnostic test is to classify or predict the presence or absence of a disease. The clinical performance of a diagnostic test is based on its ability to correctly classify subjects into relevant subgroups.

Test accuracy studies often evaluate a single index test but where alternative tests exist that can be used at the same point in the diagnostic pathway (providing the tests do not interfere with each other and the patient burden is not too great), these tests can be evaluated in one study population (Takwoingi *et al* 2013) [6]

Valid estimates of comparative accuracy can be obtained from comparative diagnostic test accuracy studies: a comparison of accuracy of two or more tests (none of which is the reference standard) within a single study (Leeflang &, Reitsma 2018) [7].

3.1 Sample Size

Sample size calculations for test accuracy studies should be determined prior to recruitment (Obuchowski 1998 & Pepe 2003) [14, 15]. When evaluating a single test, a common approach is based on the precision around an estimate of diagnostic tests. The precision of the sensitivity estimate will increase with the number of participants with the target condition (reference standard positive) and the precision of the specificity estimate will increase with the number of participants without the target condition (reference standard negative). Hence, it is vital to have an estimate of the prevalence of the target condition in the study population to plan the sample size. Adequate sample size is also critical; methods for determining sample size for diagnostic accuracy studies (Obuchowski 1998) [14]

3.2 Study Design

Study design considerations for diagnostic test accuracy studies including prospective and retrospective designs. Results of such studies are subject to several potential biases some of these biases can be avoided by careful study design, and some can be corrected to some degree in the analysis (Begg 1991) [16] identifies the most prominent and important biases as those concerned with issues relating to the reference test, or `gold standard', used to determine true disease status. Verification bias in the reported sensitivity and specificity estimates occurs if the selection of patients to receive the reference test is influenced by the result of the diagnostic test being studied (Shapiro 1999) [29].

A prospective study design in which all subjects in a defined cohort receive the reference test is preferred, but selective verification may be unavoidable for some diseases, such as when the reference test is invasive; methods for correction of verification bias exist and have been reviewed recently by Zhou (Zhou 1998) [9].

In some diagnostic accuracy studies, the test results of a series of patients with an established diagnosis are compared with those of a control group. Such case– control designs are intuitively appealing, but they have also been criticized for leading to inflated estimates of accuracy (Rutjes *et al* 2005).

3.3 Data Collection

A survey of renal failure patients was conduct to obtain the data which used for analysis to classify the new patient entries to the study. Data include only those RF patients who are on dialysis or have received a kidney transplant (i.e., treated RF). Questionnaire was use to get a data, a 94 patients give him a questionnaire whom treated in ALSELAH ALTIBI MEDICAL COMPLEX, the follow-up study was use as methodology to offer the data.

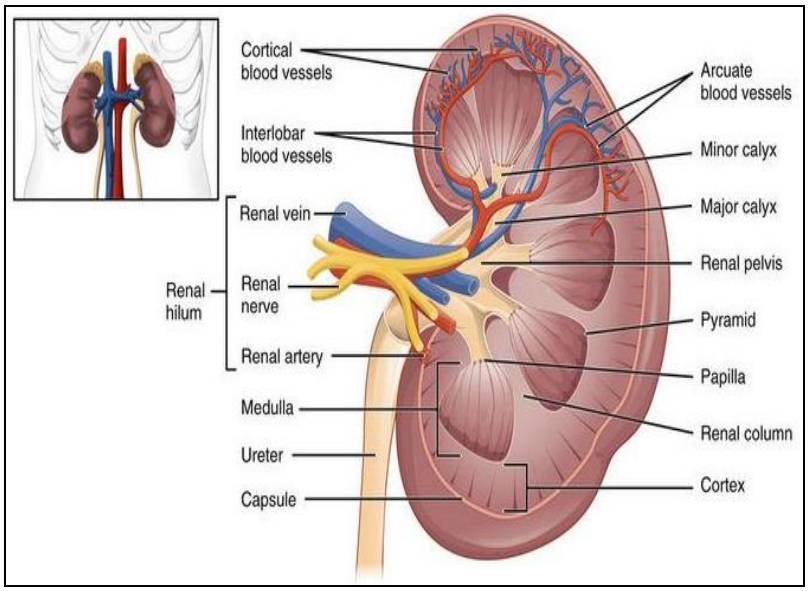
4. Chronic Kidney Disease Symptoms and Causes

Chronic kidney disease (CKD) has numerous causes. The most common causes of CKD are diabetes mellitus and long-term, uncontrolled hypertension. Polycystic kidney disease is another well-known cause of CKD (Johnson *et al* 2019) [3].

Drug overdoses, accidental or from chemical overloads of drugs such as antibiotics or chemotherapy, may also cause the onset of acute kidney injury. Unlike chronic kidney disease, however, the kidneys can often recover from acute kidney injury, allowing the patient to resume a normal life. People suffering from acute kidney injury require supportive

treatment until their kidneys recovers function, and they often remain at increased risk of developing future kidney failure (PAHO 2020) [4]. Overuse of common drugs such as ibuprofen, and acetaminophen (paracetamol) can also cause chronic kidney damage (Silver 2017).

Among the accidental causes of renal failure is the crush syndrome, when large amounts of toxins are suddenly released in the blood circulation after a long compressed limb is suddenly relieved from the pressure obstructing the blood flow through its tissues, causing ischemia. The resulting overload can lead to the clogging and the destruction of the kidneys (Bikbov 2020) [2].



https://www.mayoclinic.org/diseases-conditions/kidney.../symptoms.../syc-20369048.

Fig 1: kidney figure

5. Analysis, Discussions and Conciliations

The main confounding variables studies are gender to calculate diagnostics test for male with female because most investigators whom applicable to study male, rather than female, other confounder variable is survive patient to know how long patient live or death, the main other variable which diagnostic tests calculated were age for those younger (under age 50) or older (over age 50), time of dialysis per week to know which more likely to suspected to (CKD) whom dialysis two times per week or whom dialysis more than two times, history of disease is main variable to know if the hypertension (HTN) is more caused to (CKD) rather than diabetes mellitus (DM)), and duration of disease to study whom suspected since 5 years or more than 5 years, and mainly for those covered study in ALSELAH ALTIBI MEDICAL COMPLEX. Many epidemiological measurements calculated to get a result in case this situations such as Bayesian method and diagnostic test specially DOR, RR, Se, Sp, LR+, LR-, PPV and NPV.

Table 2: Descriptive statistics

Variables		Gender		Total
		Male	Female	
Age	Under age 50	59	28	87
	Over age 50	5	2	7
	Total	61	33	94
Time dialysis	Less than 2	13	5	18
	2 and More	48	28	76
	Total	61	33	94
History of disease	HTN	53	5	58
	DM	8	28	36
	Total	61	33	94
Duration of disease	Less than 5	46	3	49
	More than 5	15	30	45
	Total	61	33	94

Table 3: Descriptive Statistics

Variables		Survive		Total
		Life	Death	
Age	Under age 50	48	1	49
	Over age 50	42	3	45
	Total	90	4	94
Time dialysis	Less than 2 times	16	2	18
	More than 2 times	74	2	76
	Total	90	4	94
History of disease	HTN	55	3	58
	DM	35	1	36
	Total	90	4	94
Duration of disease	Less than 5	48	1	49
	More than 5	42	3	45
	Total	90	4	94

Table 4: Diagnostic tests for the variables respect to gender

Epidemiology measurements	Gender			
	Age	time dialysis per week	History of disease	Duration of disease
DOR	4.76	1.50	37.1	30.7
RR	1.14	1.40	5.79	8.33
Sens	0.68	0.72	0.91	0.94
Spec	0.71	0.37	0.78	0.67
PPV	0.97	0.21	0.13	0.75
NPV	0.15	0.85	0.85	0.91
J	-0.39	-0.09	-0.69	-0.61
LR+	2.30	1.10	4.10	2.80
LR-	0.45	1.30	8.70	11.2
FPR	0.29	0.63	0.22	0.33
FNR	0.32	0.28	0.09	0.06

Table 5: Diagnostic tests for the variables respect to survive

Epidemiology measurements	Survive			
	Age	time dialysis per week	History of disease	Duration of disease
DOR	3.40	0.22	0.52	3.40
RR	2.13	0.36	0.81	2.13
Sens	0.98	0.89	0.95	0.98
Spec	0.07	0.02	0.03	0.07
PPV	0.53	0.17	0.61	0.53
NPV	0.75	0.50	0.25	0.75
J	-0.05	-0.61	0.02	-0.05
LR+	1.10	0.91	0.98	1.10
LR-	3.50	0.18	0.60	3.50
FPR	0.93	0.98	0.97	0.93
FNR	0.02	0.11	0.05	0.02

Results and Recommendations

To know the effect of two main variables (gender and survive) on the study variables age (under 50 years & over 50 years), Duration of disease (less than 5 times & more than 5 times), time of dialysis per week (less than 2 per week & more than 2 per week), history of disease (hypertension (HYN) & diabetes mellitus (DM)). A 94 patient studied according to the variables using a questionnaire method pike-out of records in ALSILAH ALTIBI MEDICAL COMPLEX and main diagnostic test calculated rather than other measures related, for gender DOR was 4.76 (age), 1.5 (time dialysis) 37.1 (history of disease), and 30.7 (duration of disease). RR was 1.14 (age), 1.4 (time dialysis) 5.79 (history of disease), and 8.33 (duration of disease). Sensitivity was 0.68 (age), 0.72 (time dialysis) 0.91 (history of disease, and 0.94 (duration of disease). Specificity was 0.71 (age), 0.37 (time dialysis) 0.78 (history of disease), and 0.67 (duration of disease). PPV was 0.97 (age), 0.21 (time dialysis) 0.13 (history of disease, and 0.75 (duration of disease). NPV was 0.15 (age), 0.85 (time dialysis) 0.85 (history of disease, and 0.91 (duration of disease). LR+ was 2.30 (age), 1.10 (time dialysis) 4.10 (history of disease, and 2.80 (duration of disease). LR- was 0.45 (age), 1.30 (time dialysis) 8.70 (history of disease, and 11.20 (duration of disease). Youden index (J) was -0.39 (age), -0.09 (time dialysis) -0.69 (history of disease, and -0.61 (duration of disease).

In the general, most diagnostic test calculated of chronic kidney disease risk factors was significantly associated among the people studied a person was the male is more likely in the older rather younger, and in those whom two times dialysis per week than whom more than two times, and in history of disease with those suspected of (HTN) more likely than (DM), and on duration of disease for those duration more than 5 years than whom less than 5 years.

With regard to survive, DOR was 3.40 (age), 0.22 (time dialysis) 0.52 (history of disease), and 3.40 (duration of disease). RR was 2.13 (age), 0.36 (time dialysis) 0.81 (history of disease), and 2.13 (duration of disease). Sensitivity was 0.98 (age), 0.89 (time dialysis) 0.95 (history of disease, and 0.98 (duration of disease). Specificity was 0.07 (age), 0.02 (time dialysis) 0.03 (history of disease), and 0.07 (duration of disease). PPV was 0.53 (age), 0.17 (time dialysis) 0.61 (history of disease, and 0.53 (duration of disease). NPV was 0.75 (age), 0.50 (time dialysis) 0.25 (history of disease, and 0.75 (duration of disease). LR+ was 1.10 (age), 0.91 (time dialysis) 0.98 (history of disease, and 1.10 (duration of disease). LR- was 3.50 (age), 0.18 (time dialysis) 0.60 (history of disease, and 3.50 (duration of disease). Youden index (J) was -0.05 (age), -0.61 (time dialysis) 0.02 (history of disease, and -0.05 (duration of disease).

In the general, most diagnostic test calculated of chronic kidney disease risk factors was significantly associated among the people studied a person was the life is more likely in the younger rather older, and in those whom two times dialysis per week than whom more than two times, and in history of disease with those suspected of (HTN) more likely than (DM), and on duration of disease for those duration less than 5 years than whom more than 5 years.

Conclusion

Screening tests, it is important to avoid misconceptions about diagnostic tests. Clarification is then provided about the definitions of them and why investigators and clinicians can misunderstand and misrepresent them. Reasoning are made that main diagnostic test should usually be applied only in the context of describing a screening test's attributes relative to a reference standard. The most commonly used measures of accuracy or validity are sensitivity and specificity. Sensitivity refers to its ability to detect a high proportion of the true cases, that is, to yield few false negative results. Specificity refers to the fact that a specific test is the one that correctly identifies the true negative, and hence yields few false positive verdicts (Brenner 1997). The whole point of diagnostic test is to use it to make a diagnosis, so we need to know the probability that the test will give the correct diagnosis. The sensitivity and specificity do not give us this information. Instead, we must approach the data from the direction of the test results, using predictive values. The probability of disease, given the result of a test is called the predictive value of the test (Altman 1994) [10]. Positive predictive value (PPV) is the probability that a subject with a positive screening test is actually suffering from the disease. Negative predictive value (NPV) is the probability that a subject diagnosed by a negative screening test result truly do not have the disease. Hence, both the predictive values are associated with the diagnosis being correct (Fletcher 2014). However; predictive value of the test depends on prevalence of the disease. Another useful measure of the accuracy of a diagnostic test is the likelihood ratio (LR). LR indicates the value of the test for increasing certainty about a positive diagnosis. For any test results we can compare the probability of getting that result if the patient truly had the condition of interest with the corresponding probability of s/he were healthy. The ratio of these two probabilities is called the LR (Dhamnetiya *et al* 2022) [20]. Diagnostic Odds Ratio is also one global measure for diagnostic accuracy, used for general estimation of discriminative power of diagnostic procedures and also for the comparison of diagnostic accuracies between two or more diagnostic tests (Glas *et al* 2003) [19]. Another index named as Youden's index is one of the oldest measures for diagnostic accuracy. It is also a global measure of a test performance, used for the evaluation of overall discriminative power of a diagnostic procedure (youden 1950) [12]. A large DOR may be explained by a high sensitivity combined with a poor specificity or vice versa. In general, it is preferred to describe the accuracy of a test using both sensitivity and specificity.

References

1. Arias E, Anderson RN, Kung H-C, Murphy SL, Kochanek KD. Deaths: final data for 2001. National vital statistics reports 52 no. 3. Hyattsville, Maryland: National Center for Health Statistics,2003:9:30-3.
2. Bikbov B, Purcell CA, Levey AS, *et al*. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the global burden of disease study. Lancet,2020:395:709-733, 2017.
3. Johnson RJ, Wesseling C, Newman LS. Chronic kidney disease of unknown cause in agricultural communities. N Engl J Med.,2019:380(19):1843-1852.
4. PAHO Progress reports on Technical Matters. Chronic kidney disease in agricultural communities in Central America. CE156/INF/8. 156th Session of the executive committee, Washington, D.C., USA, 2020, 17-21. Available at: https://www.paho.org/hq/dmdocuments/2015/CE156-INF-8-e.pdf. Accessed January 6.
5. Silver SA, Long J, Zheng Y, Chertow GM. Cost of acute kidney injury in hospitalized patients. J Hosp Med.,2017:12(2):70-6.
6. Takwoingi Y, Leeflang MMG, Deeks JJ. Empirical Evidence of the importance of comparative studies of diagnostic test accuracy. Ann Intern Med,2013:158(7):544. doi: 10.7326/ 0003- 4819- 158-7- 201304020- 00006.
7. Leeflang MMG, Reitsma JB. Systematic reviews and meta-analyses addressing comparative test accuracy questions. Diagnostic Progn Res,2018:2(1):17. doi: 10.1186/s41512- 018- 0039- 0.
8. Newcomer SR, Xu S, Kulldorff M, Daley MF, Fireman B, Glanz JM. A primer on quantitative bias analysis with positive predictive values in research using electronic health data. J Am Med Inform Assoc,2019:26(12):1664-74. doi: 10.1093/ jamia/ ocz094.
9. Zhou XH, Obuchowski NA, Obuchowski DM. Statistical Methods in Diagnostic Medicine. New York: John Wiley and Sons, 2002.
10. Altman DG, Bland JM: Diagnostic tests. 1. Sensitivity and specificity. BMJ,1994:308:1552.
11. Deeks JJ, Altman DG: Diagnostic tests. 4. Likelihood ratios. BMJ,2004:329:168.
12. Youden WJ. (Index for rating diagnostic tests. Cancer,1950:3(1):32-35.
13. Werner Vach, Veronika Reiser, Izabela Kolankowska, Susanne Weber, Gerta Rücker, Design and Evaluation of Diagnostic Studies. Freiburg, 2017.
14. Obuchowski NA. Sample size calculations in studies of test accuracy. Statistical Methods in Medical Research,1998:7:371-392. (https://doi.org/10.1177/096228029800700405)
15. Pepe MS, Pepe PBMS. The Statistical Evaluation of Medical Tests for Classification and Prediction. Oxford University Press, 2003.
16. Begg CB. Advances in statistical methodology for diagnostic medicine in the 1980's. Statistics in Medicine,1991:10:1887- 95.
17. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. Stat Med,1997:16(09):981-991.
18. Fletcher RH, Fletcher SW, Fletcher GS. Diagnosis. Clinical Epidemiology: The Essentials. 5th ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins, 2014, 108-131.
19. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol,2003:56(11):1129-1135.
20. Deepak Dhamnetiya, Ravi Prakash Jha, Shalini, Krittika Bhattacharyya, How to Analyze the Diagnostic Performance of a New Test? Explained with Illustrations. The Indian Association of Laboratory Physicians,2022:14:90-98.
21. Eusebi P. Diagnostic accuracy measures. Cerebrovasc, Dis,2013:36(4):267-72.
22. Abdul Ghaaliq Lalkhen, Anthony McCluskey Clinical tests: sensitivity and specificity Continuing Education in Anaesthesia, Critical Care & Pain,2008:8(6):221-223.
23. ˇSimundić A-M. Measures of diagnostic accuracy: basic definitions. Ejifcc.,2012:19(4):203. PMID: 27683318.
24. Naaktgeboren CA, De Groot JAH, Rutjes AWS, Bossuyt PMM, Reitsma JB, Moons KGM. Anticipating missing reference standard data when planning diagnostic accuracy studies. BMJ (Online), 2016, 352. https://doi.org/10.1136/bmj.i402 PMID: 26861453.
25. Thompson M, Van den Bruel A. Sources of Bias in Diagnostic Studies. Diagnostic Tests Toolkit: Wiley-Blackwell,2011, 26-33.
26. Xue X, Oktay M, Goswami S, Kim MY. A method to compare the performance of two molecular diagnostic tools in the absence of a gold standard. Statistical Methods in Medical Research,2019:28(2):419-31.
27. Thomas R Vetter, Patrick Schober, MM ed Stat, Edward J Mascha. Diagnostic Testing and Decision-Making: Beauty Is Not Just in the Eye of the Beholder SPECIAL ARTICLE,2018:127:1085-91.
28. Bolin E, Lam W. A review of sensitivity, specificity, and likelihood ratios: evaluating the utility of the electrocardiogram as a screening tool in hypertrophic cardiomyopathy. Congenit Heart Dis. Sep-Oct,2013:8(5):406-10.
29. Shapiro DE. The interpretation of a diagnostic tests, Statistical Methods in Medical Research,1999:8:113-134.
30. Sergeant ESG. Epitools epidemiological calculators. Ausvet, 2018. Available at: http://epitools.ausvet.com.auAccessed January 15, 2021.
31. Alice J Sitch, Olaf M Dekkers, Barnaby R Scholefield, Yemisi Takwoingi, Introduction to diagnostic test accuracy studies, Methodology Editorial European Journal of Endocrinology,2021:184:E5-E9.