REVIEW ARTICLE

Risk of Bias Assessment and Risk Minimisation Strategies in COVID-19 Diagnostic Test Accuracy Study

Muhammad Ikhwan Mud Shukri, Nur Za'Imah Zainal, Ahmad Zaid Fattah Azman

Department of Community Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

ABSTRACT

It is paramount to assess the risk of biases in may arise from diagnostic test accuracy (DTA) study as it will affect the accuracy and validity of the tests. These biases can be found in published researches and here we look at COVID-19 DTA studies. The evaluation of bias risk in diagnostic research is mainly performed using QUADAS-2. The aim of this review was to determine potential selection and information biases in diagnostic test accuracy studies and strategies to minimize risk of biases. Literature review related to diagnostic test accuracy study is identified through an online search of databases namely PubMed, ScienceDirect, Research Gate, Google Scholar, and official government websites range. Six potential biases in four QUADAS-2 domains are identified in COVID-19 diagnostic test accuracy study which are 1) spectrum bias in patient selection; 2) interpretation bias in index test; 3) differential misclassification bias and nondifferential misclassification bias in reference standard; and 4) partial verification bias and differential verification bias in patient flow. The identified biases exert effects on accuracy of COVID-19 diagnostic tests. Six strategies are recommended to reduce these biases, hence, improving the accuracy of COVID-19 diagnostic tests. The best diagnostic test can give benefits to the population in the mass screening program during COVID-19.
Malaysian Journal of Medicine and Health Sciences (2024) 20(1):359-364. doi:10.47836/mjmhs.20.1.43

Keywords: Bias, diagnostic test accuracy study, QUADAS-2, COVID-19

Corresponding Author:
Ahmad Zaid Fattah Azman, DrPH
Email: azfa@upm.edu.my
Tel: +603-97692943

INTRODUCTION

The diagnostic test accuracy (DTA) study development for both clinical and epidemiological purposes is a critical step in the COVID-19 pandemic response(1). It guides researchers in establishing sound methodology of conducting DTA study to ensure validity and reliability of COVID-19 tests (2). In the public health setting, the DTA studies are used in identifying screening tests for the benefit of the community (3). The inaccuracy results of DTA study exert adverse complications such as wrong identification of COVID cases. Wrong case identification leads to higher case mortality, ineffective COVID-19 control due to failure to isolate false-negative cases and inefficient utilization of resources due to unnecessary false-positive case treatment (4). Therefore, DTA study should be prioritised to complement existing COVID-19 control program. The main components of the DTA study are the index and reference tests (5). For instance, to conduct a DTA study for COVID-19 test, the index test is a rapid saliva test and the reference test is the nasopharyngeal swab test for PCR (6). Since the Covid-19 pandemic hit the world, there are many DTA studies conducted to find accurate tests to detect SARS-COV-2 (7). The pandemic gives a big impact on the world, especially the researchers. Study design issues that may skew estimates of test accuracy are becoming more widely recognised (8).

The definition of bias in diagnostic studies is a systematic deviation from the true value in an observed measurement (9). Biases will arise in diagnostic studies if they are not controlled properly. Bias develops in DTA study when the values of sensitivity or specificity differs from the actual value. It also can be defined as any systematic inaccuracy in a study's design, execution, or analysis that leads to an incorrect assessment of an exposure's impact on the risk of disease (10,11). For instance, bias can arise from the initial part of the study which is from selecting the patient, the interpretation of index test, and references test and the patient flow.

There are several tools in identifying biases and assisting in reducing biases. These tools assist researchers to report and appraise DTA study results (8,12). An example of a quality assessment tool that can be used to assess risk of bias in COVID-19 DTA study is QUADAS-2. QUADAS-2 is the latest version of QUADAS (Quality Assessment of Diagnostic Accuracy Studies), which primarily aim to assess the risk of bias that arises from methodology part, which are the selection of patient,

data collection, test execution or interpretation, or data analysis (11,13,14). The QUADAS-2 was developed based on the QUADAS tool which was created in 2003 (14).The two main components of QUADAS-2, are the risk of bias and concerns about applicability (13,14). The QUADAS-2 tool is divided into four domains: patient selection, index test, reference test, and patient flow. The QUADAS-2 format is very similar to the PICO (patient or population/intervention/comparator/outcomes) format (15). All key domains are scored based on bias risk and their appropriateness to research questions. Each key domain has a set of signalling questions to aid in bias and applicability judgments. The concern about applicability was included in the first three domains as well as evaluation of term of bias risk in each domain (16). Before entering the four domains in the QUADAS-2 checklist, the research question should be clearly stated with illustration of a flow diagram for the primary study (8,13).

The main objective of this review was to determine the potential biases that can arise during COVID-19 DTA studies based on QUADAS-2 tool and strategies to minimize risk of biases. This information should assist public health physicians and clinical specialists in identifying bias risks in COVID-19 DTA studies, anticipating the effects of bias on researches, and avoiding methodologic issues resulted from biases in DTA studies.

METHODOLOGY

A literature review was conducted from April 2022 to June 2022 by searching online databases, namely PubMed, ScienceDirect, Research Gate, Google Scholar, and official government websites on the risk of bias assessment in COVID-19 diagnostic test accuracy study. The search terms used were as follows: "diagnostic test accuracy study", "bias AND QUADAS-2", "risk of bias AND diagnostic test accuracy study" and "bias AND diagnostic study COVID-19". The selected information used in this review is taken from any journal, article, report, or news update regarding the topic. Inclusion criteria for selection were articles written in English and published between 1st January 2012 and 31st March 2022.

RESULTS AND DISCUSSION

Based on QUADAS-2, six types of biases were identified to be potentially arisen during COVID-19 DTA study. The six biases can be largely classified into two major types of biases which are selection biases (comprised of spectrum bias, partial verification bias and differential verification bias) and information biases (comprised of interpretation bias, differential misclassification bias and nondifferential misclassification bias). Based on QUADAS-2 domains, the mechanism of these biases occurrence and strategies to minimize the risk of biases were discussed.

Risk Of Bias Based On Quadas-2 Domains And Risk Minimisation Strategies

Domain 1: Patient Selection

Patient selection is the selection of participant from a population of interest (15). The process of selecting the target population depends on clinical questions or PICO elements (17). One of the key components in the selection of the participants is internal validity (18). The internal validity of clinical research largely depend on study design, data collection, and statistical analyses, and it is threatened by biases and random variation (19). One of the biases identified in COVID-19 DTA study is spectrum bias. There no definitive definition of spectrum bias, however, it can generally be defined as bias resulting from changes of tests outcomes when applied in different populations (20). This bias is produced when researchers included only ''clear'' or ''definite'' cases, not representing the whole spectrum of disease, and/or ''clear'' or healthy control subjects, not representing the conditions in which a differential diagnosis should be carried out. A particular case is the purity diagnostic bias, when selecting cases of a certain disease those with other comorbidities are excluded and the final sample does not represent the cases that originated (8). For example, during COVID-19 pandemic, serological tests were highly in demand to determine the COVID-19 serological status to address a variety of needs, such as determining COVID-19 infection rate, case fatality rate, virus load and transmission at the population level. The spectrum bias appeared during the early period of the test development, which is from establishment test testing to population-based testing (21).

The process of illness that ranges from mild to severe or fatal in individuals is referred to as the disease spectrum. In Malaysia, for example, COVID-19 patients are classified into five levels or clinical categories, consists of asymptomatic, symptomatic without pneumonia, symptomatic with pneumonia, symptomatic with pneumonia and requiring oxygen assistance, and critically ill patients with multiple organ complications (22). Spectrum bias may happen when COVID-19 spectrum deviates from the clinical categories or ambiguous results were excluded from clinical categories. It may also occur in COVID-19 case control studies. In this study design, an index test was used in two groups, one for the case and one for the control. Researchers may introduce spectrum bias by failing to include certain group of patients who represent the certain disease spectrum in case control study (23). Therefore, to minimize spectrum bias in COVID-19 DTA study, it is recommended to consider changing COVID-19 DTA study design from case control to cohort design to include wide spectrums of patients.

Domain 2: Index Test

Index test is defined as the diagnostic test evaluated in comparison to reference test (22). For example, COVID-19 infection is diagnosed by identifying viral RNA in samples collected via a nasopharyngeal swab or other respiratory samples. However, it has several limitations for use in mass screening, the most important of which are the time required for diagnosis, the overcrowding of those centres designated to analyse the specimens, and the non-negligible risk of viral transmission to healthcare workers (24). Using saliva as a diagnostic sample has advantages such as being easily provided by the patient and being collected without the need for specialized staff. As a result, the Rapid Saliva Test (RST) should be used as a diagnostic test for a mass screening programme (6).

However, there are risk of bias in performing the index test including bias during conducting the test and bias from the interpretation of the result from diagnostic test. The interpretation of diagnostic accuracy study is a must, and the way of the interpretation can affect the test's performance. The interpretation of the index test relies on the competency of the interpreter or the index test flaws. The interpretation bias can arise because of indeterminate result or review of bias. Due to indeterminate results, it can give rise to interpretation bias. Indeterminate results may give rise to spectrum bias if they are excluded from the analysis. However, researchers should be critical in analysing indeterminate results to be positive or negative if indeterminate results are not excluded. For instance, a very striking feature was observed when comparing the results of the salivary rRT-PCR with those of the nasopharyngeal swab in the subjects who had been previously classified as false negatives and false positives with the RST. The two subjects who were classified as false negatives tested also negative by salivary rRT-PCR, thus the viral RNA was not detected in the saliva (6,25).

For interpretation bias may occur in diagnostic test studies when there is no blinding of gold standard (verification test) or its result (24). This is because the majority of clinicians who interpret tests are influenced by prior knowledge or the available results of index tests. According to the international guidelines, the nasopharyngeal swab was analysed by independent blinded clinicians using real-time reverse transcription (rRT)-PCR. The salivary sample collected for the RST was examined by rRT-PCR to provide data on the presence of the virus in the saliva and to better analyse any discrepancy between the RST and nasopharyngeal swab results (6). Therefore, to minimize the risk of interpretation bias in COVID-19 DTA study, blinding of clinicians should be carried. If a clinician is a researcher, double blinding should be used whereas if a clinician is in the team analysing the data, triple blinding should be used.

Domain 3: Reference Test

Reference standard is defined as a test, combination of tests or procedure that is considered the best available method to categorize participants in diagnostic test accuracy study into either group of having a condition or disease or group free of that condition or disease (13). In DTA study, it is sometimes synonymous to reference test or gold standard. It is usually utilized in comparing index test accuracy in accordance to reference standard. For example, the saliva nucleic acid amplification testing (NAAT) used as rapid COVID-19 test kit serves as index test whereas nasopharyngeal RTK Ag and RT-PCR serve as a reference tests (26).

In reference standard, classification bias can happen due to systematic error. Classification bias can be differential misclassification and nondifferential misclassification (27). Differential misclassification is related to exposure, outcome or treatment whereas nondifferential misclassification occurs at random (28). In COVID-19 testing, differential misclassification may happen during study design stage based on the assumption that untested individuals are uninfected. Due to this assumption, infection rate will be higher among untested individuals as compared to tested individuals (29). Differential misclassification occurred as a result from increase or decrease of risk factors of COVID-19 infection such as age and gender (30). For example, in some countries or health facilities, elder peoples aged 60 years are routinely screened for COVID-19 infection. Elderly people will be more likely to be tested, therefore, less likely to be misclassified as uninfected due to not receiving testing. To minimize this bias, it is recommended to conduct COVID-19 DTA study in large sample size.

Nondifferential misclassification may happen when exposed or true positive COVID-19 person is incorrectly classified into non-exposed or negative COVID-19 group. In COVID-19 DTA study, an example of nondifferential misclassification and its effects on sensitivity and specificity of index test is illustrated in Table I based on hypothetical results. Based on Table I, in reference test, the true positive number is lowered by 10% nondifferential misclassification rate, thus affecting its accuracy. To reduce differential misclassification, the accuracy of COVID-19 index test should be compared to the most accurate reference test available (31).

Domain 4: Patient flow

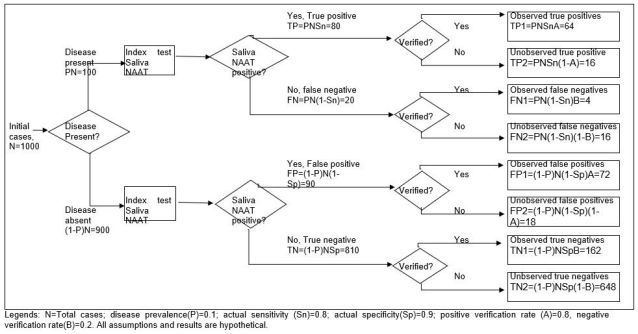
In this domain, verification bias may occur as a result of diagnosis of disease among studied samples not receiving same reference standard (32). It can be classified into partial or differential verification bias (32). Partial verification occurs when studied samples tested with index test failed to do reference tests. Failure to do verification may result from the preliminary result being negative or the refusal to do reference tests (33). In COVID-19 DTA study, for example, nasopharyngeal

Table I: Nondifferential misclassification and its effect on sensitivity and specificity of COVID-19 Saliva test

	(Perfect Reference Test)		(Imperfect Reference Test)	
	Reference Test 1		Reference Test 2 (10% Misclassification)	
	RT-PCR Test		Nasopharyngeal RTK Ag Test	
Saliva NAAT (Index Test)	Positive	Negative	Positive	Negative
Positive	950	50	860*	140
Negative	50	950	140	860
Total	**1000**	**1000**	**1000**	**1000**
Sensitivity	0.95		0.86	
Specificity	0.95		0.86	

*True positive= 950(1-0.1) + 50(0.1) **Results are hypothetical.

NAAT using PCR test is considered as gold standard. However, the test is expensive or painful during nasopharyngeal sampling, which may cause those who are tested positive saliva test to default reference test for verification (34). Figure 1 demonstrated the partial verification bias effect. Among 1000 people with assumption of prevalence of COVID-19 of 10%and hypothetical saliva NAAT sensitivity of 80% and specificity of 90%, the positive cases are verified at higher rate (80%) than negative cases (20%). The results from Figure 1 are summarized in Table II. Table II showed the observed accuracy in comparison with actual accuracy (accuracy obtained with full verification). The observed accuracy demonstrated false increase of sensitivity from 0.80 to 0.94 and false decrease of specificity from 0.90 to 0.6. Therefore, partial verification bias may increase COVID-19 index test sensitivity and reduce its specificity. In this case, the verification rate is assumed to depend on the index test result. Partial verification bias, however, is impossible where positive and negative verification rate is identical. Thus, COVID-19 full verification test is recommended to eliminate partial verification bias with practical consideration on costs and resources.



Legends: N=Total cases; disease prevalence(P)=0.1; actual sensitivity (Sn)=0.8; actual specificity(Sp)=0.9; positive verification rate (A)=0.8, negative verification rate(B)=0.2. All assumptions and results are hypothetical.

Figure 1: Flow diagram for partial verification bias.

On the other hand, differential verification bias occurs when index case is verified by multiple reference standard test (25). Multiple reference tests are used due to test invasiveness, high costs (32) or different outcome result (33). In some countries such as Malaysia, nasopharyngeal antigen test kit and RT-PCR can be used as diagnostic tests (26). However, the practice of using two or more standard tests to verify COVID-19 leads to differed accuracy. The accuracy of NP antigen test kit and RT-PCR may differ and the accuracy of Saliva NAAT used to refer for verification can result into classification bias. The overall accuracy of unbiased and biased result will result in differential verification bias.

Table III showed the differential verification bias effect on the accuracy of overall tests. In the previous example, there are cases which are not verified (unobserved cases) with reference test (RT-PCR). The unobserved cases may be verified by other reference standard (RTK Ag) (35). Assuming the reference standard has nondifferential misclassification rate of 10%, the total accuracy of combined reference tests changed in which the sensitivity has reduced and specificity increased in comparison to Gold Standard. Therefore, to reduce differential verification bias, the flow of COVID-19 diagnostic testing should be well-designed to avoid unnecessary multiple reference tests (22).

Table II: Partial verification bias and its effect on sensitivity and specificity of COVID-19 Saliva test

	Actual		Observed	
	RT-PCR*		RT-PCR*	
Saliva NAAT**	Disease	No Disease	Disease	No Disease
Positive	80	90	64	72
Negative	20	810	4	162
Total	**100**	**900**	**68**	**234**
Sensitivity	0.80		0.94	
Specificity	0.90		0.69	

*Reference test **Index test

Table III: Differential verification bias and its effect

	Observed Cases (NP RT-PCR)*		Unobserved Cases		Observed Cases (NP RTK Ag)**		Combined (NP RT-PCR and RTK Ag)	
	Reference Test		Reference Test		Reference Test		Reference Test	
Saliva NAAT ***	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative
Positive	64	72	16	18	16.2	17.8	80.2	89.8
Negative	4	162	16	648	79.2	584.8	83.2	746.8
Total	**68**	**234**	**32**	**666**	**95.4**	**602.6**	**163.4**	**836.6**
Sensitivity	0.94		0.50		0.16		0.49	
Specificity	0.69		0.97		0.91		0.89	

*Reference Test 1(Gold standard) **Reference Test 2 ***Index Test. Assumption of Reference Test 2 has 10% nondifferential misclassification rate.

CONCLUSION

This article summarises the six types of biases that are unique to DTA study related to COVID-19 test. The objective of this review is to assist relevant stakeholders of COVID-19 DTA studies in critically appraising results with comprehensive consideration of type of biases. It is imperative to understand that no study is bias-free. Recognizing and minimizing biases through recommended strategies in COVID-19 DTA study, however, may improve the accuracy of COVID-19 diagnostic tests. Reporting the biases identified in DTA study may enhance the overall quality of the study reporting and for readers to make informed and fair judgement on its quality.

ACKNOWLEDGEMENT

The authors wish to acknowledge support from Department of Community Health, Faculty of Medicine and Health Science, Universiti Putra Malaysia for the assistance given during the review.

REFERENCES

1. Peeling RW, Heymann DL, Teo YY, Garcia PJ. Diagnostics for COVID-19: moving from pandemic response to control. Lancet. 2022;399(10326):757-768. doi:10.1016/S0140-6736(21)02346-1
2. Rohde J, Himmel W, Hofinger C, Lвm T-T, Schrader H, Wallstabe J, et al. Diagnostic accuracy and feasibility of a rapid SARS-CoV-2 antigen test in general practice – a prospective multicenter validation and implementation study. BMC Prim Care [Internet]. 2022;23(1):149. doi: 10.1186/s12875-022-01756-1
3. Takwoingi Y, Quinn TJ. Review of Diagnostic Test Accuracy (DTA) studies in older people. Age Ageing [Internet]. 2018 May 1;47(3):349–55. doi:10.1093/ageing/afy023
4. OECD. Testing for COVID-19: A way to lift confinement restrictions. Oecd [Internet]. 2020;(5):1–21. Available from: https://www.oecd.org/coronavirus/policy-responses/testing-for-covid-19-a-way-to-lift-confinement-restrictions-89756248/
5. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Clin Biochem. 2003;36(1):2-7. doi:10.1016/s0009-9120(02)00443-5
6. Azzi L, Baj A, Alberio T, Lualdi M, Veronesi G, Carcano G, et al. Rapid Salivary Test suitable for a mass screening program to detect SARS-CoV-2: A diagnostic accuracy study. J Infect. 2020;81(3):e75–8. doi: 10.1016/j.jinf.2020.06.042.
7. Spick M, Lewis HM, Wilde MJ, Hopley C, Huggett J, Bailey MJ. Systematic review with meta-analysis of diagnostic test accuracy for COVID-19 by mass spectrometry. Metabolism. 2022 Jan;126:154922. doi: 10.1016/j.metabol.2021.154922.
8. Schmidt RL, Factor RE. Understanding sources of bias in diagnostic accuracy studies. Arch Pathol Lab Med. 2013;137(4):558–65. doi: 10.5858/arpa.2012-0198-RA.
9. Mower WR. Evaluating Bias and Variability in Diagnostic Test Reports. Ann Emerg Med [Internet]. 1999;33(1):85–91. doi: 10.1016/s0196-0644(99)70422-1.
10. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995;273(5):408-412. doi:10.1001/jama.273.5.408
11. Whiting P, Rutjes AWS, Dinnes J, Reitsma JB, Bossuyt PMM, Kleijnen J. A systematic review finds that diagnostic reviews fail to incorporate quality despite available tools. J Clin Epidemiol. 2005;58(1):1–12. doi: 10.1016/j.jclinepi.2004.04.008.
12. Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ. 2006;174(4):469-476. doi:10.1503/cmaj.050090
13. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. BMJ Open. 2016;6(11):1–17. doi: 10.1136/bmjopen-2016-012799.
14. Reitsma JB, Leeflang MMG, Sterne JAC, Bossuyt PMM, Whiting PF, Rutjes AWSS, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(4):529–36. doi: 10.7326/0003-4819-155-8-201110180-00009.
15. Martínez-Mesa J, González-Chica DA, Duquia RP, Bonamigo RR, Bastos JL. Sampling: how to select participants in my research study? An Bras Dermatol. 2016;91(3):326–30. doi: 10.1590/abd1806-4841.20165254.
16. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ [Internet]. 2015;351:h5527. doi: 10.1136/bmj.h5527.
17. Aslam S, Emmanuel P. Formulating a researchable question: A critical step for facilitating good clinical research. Indian J Sex Transm Dis AIDS. 2010 Jan;31(1):47–50. doi: 10.4103/0253-7184.69003.
18. Akobeng AK. Assessing the validity of clinical trials. J Pediatr Gastroenterol Nutr. 2008 Sep;47(3):277–82. doi: 10.1097/MPG.0b013e31816c749f.
19. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection Bias and Information Bias in Clinical Research. Nephron Clin Pract [Internet]. 2010;115(2):c94–9. doi:10.1159/000312871
20. Willis BH. Spectrum bias—why clinicians need to

be cautious when applying diagnostic test studies. Fam Pract [Internet]. 2008 Oct 1;25(5):390–6. doi:10.1093/fampra/cmn051
21. Einhauser S, Peterhoff D, Niller HH, et al. Spectrum Bias and Individual Strengths of SARS-CoV-2 Serological Tests-A Population-Based Evaluation. Diagnostics (Basel). 2021;11(10):1843. doi:10.3390/diagnostics11101843
22. Ministry of Health Malaysia. Clinical Management of Confirmed COVID-19 in Adult and Paediatric. 2022;(May). Available from: https://www.cdc.gov/coronavirus/2019-ncov/science/science-
23. Hall MK, Kea B, Wang R. Recognising Bias in Studies of Diagnostic Tests Part 1: Patient Selection. Emerg Med J. 2019;36(7):431-434. doi:10.1136/emermed-2019-208446
24. Azzi L, Maurino V, Baj A, Dani M, d'Aiuto A, Fasano M, et al. Diagnostic Salivary Tests for SARS-CoV-2. J Dent Res. 2021;100(2):115–23. doi: 10.1177/0022034520969670
25. Kea B, Hall MK, Wang R. Recognising bias in studies of diagnostic tests part 2: interpreting and verifying the index test. Emerg Med J. 2019;36(8):501-505. doi:10.1136/emermed-2019-208447
26. Health M of. Annex 2h Flowchart for the Management of Probable COVID-19 Case. 2021;42. Available from: https://covid-19.moh.gov.my/garis-panduan/garis-panduan-kkm
27. Last JM. A dictionary of epidemiology. Int J Epidemiol. 1986;15(2):277. doi: 10.1093/ije/15.2.277
28. Lambert J. Statistics in brief: How to assess bias in clinical studies? Clin Orthop Relat Res. 2011;469(6):1794–6. doi: 10.1007/s11999-010-1538-7.
29. Accorsi EK, Qiu X, Rumpler E, Kennedy-Shaffer L, Kahn R, Joshi K, et al. How to detect and reduce potential sources of biases in studies of SARS-CoV-2 and COVID-19. Eur J Epidemiol [Internet]. 2021;36(2):179–96. doi: 10.1007/s10654-021-00727-7
30. Sim BLH, Chidambaram SK, Wong XC, et al. Clinical characteristics and risk factors for severe COVID-19 infections in Malaysia: A nationwide observational study. Lancet Reg Health West Pac. 2020;4:100055. doi:10.1016/j.lanwpc.2020.100055
31. Brenner H, Blettner M. Misclassification bias arising from random error in exposure measurement: implications for dual measurement strategies. Am J Epidemiol. 1993;138(6):453-461. doi:10.1093/oxfordjournals.aje.a116877
32. O'Sullivan JW, Banerjee A, Heneghan C, Pluddemann A. Verification bias. BMJ Evidence-Based Med [Internet]. 2018 Apr 1;23(2):54 –55. doi: 10.1136/bmjebm-2018-110919.
33. Pluddemann A, McCall M, O'Sullivan J BA. Verification bias - Catalog of Bias [Internet]. 2019 [cited 2022 Jan 12]. Available from: https://catalogofbias.org/biases/verification-bias/
34. Rubin R. The Challenges of Expanding Rapid Tests to Curb COVID-19. JAMA - J Am Med Assoc. 2020;324(18):1813–5. doi: 10.1001/jama.2020.21106.
35. Yland JJ, Wesselink AK, Lash TL, Fox MP. Misconceptions About the Direction of Bias From Nondifferential Misclassification [published correction appears in Am J Epidemiol. 2022 Nov 19;191(12):2123]. Am J Epidemiol. 2022;191(8):1485-1495. doi:10.1093/aje/kwac035