Systematic review of aortic dissection detection risk score plus D-dimer for diagnostic rule-out of suspected acute aortic syndromes

SAEM Academic Emergency Medicine
Official Journal of the Society for Academic Emergency Medicine

Systematic review of aortic dissection detection risk score plus D-dimer for diagnostic rule-out of suspected acute aortic syndromes

Journal:	*Academic Emergency Medicine*
Manuscript ID	AEMJ-19-995.R2
Manuscript Type:	Systematic Reviews (With or Without Meta-analyses)
Classifications:	Cardiovascular Emergencies, Diagnostic Testing (Non-Imaging)

ABSTRACT

Objectives. In patients at low clinical probability of acute aortic syndromes (AASs), decision on advanced aortic imaging is cumbersome. Integration of the aortic dissection detection risk score (ADD-RS) with D-dimer (DD) provides a potential pipeline for standardized diagnostic rule-out. We systematically reviewed and summarized supporting data.

Methods. Cross-sectional studies assessing integration of ADD-RS with DD for diagnosis of AASs were identified on MEDLINE, EMBASE and Web Of Science databases. Two reviewers independently screened articles, assessed quality and extracted data. The quality of design and reporting was evaluated with the QUADAS-2 and STARD tools. Individual patient data were obtained, to allow analysis of both conventional (500 ng/mL) and age-adjusted ($DD_{age\text{-}adj}$) DD cutoffs. Data were summarized for 4 diagnostic strategies combining ADD-RS=0 or ≤1, with DD<500 ng/mL or <$DD_{age\text{-}adj}$. The statistical heterogeneity of the diagnostic variables was estimated with Higgins' I^2. Pooled values were calculated for variables showing non-significant heterogeneity.

Results. After screening of 680 studies, 4 articles (including a total of 3804 patients) met inclusion criteria. One prospective study provided a low risk of bias/applicability concerns, while methodological limitations were found in the other 3 retrospective studies. Statistical heterogeneity was negligible for sensitivity and negative likelihood ratio (LR) values, and significant for specificity and positive LR values of all diagnostic strategies. Pooled sensitivity was 99.9% (95%CI 99.3–100%, I^2=0) for ADD-RS=0 and DD<500 ng/mL or <$DD_{age\text{-}adj}$, 98.9% (95%CI 97.9–99.9%, I^2=0) for ADD-RS≤1 and DD<500 ng/mL and 97.6% (95%CI 96.3–98.9%, I^2=0) for ADD-RS≤1 and DD<$DD_{age\text{-}adj}$.

Conclusions. Despite methodological limitations, integration of ADD-RS=0 or ≤1 with DD<500 ng/mL shows negligible heterogeneity and consistently high sensitivity across studies, thus supporting reliability for diagnostic rule-out of AASs. Data supporting ADD-RS=0 plus $DD_{age\text{-}adj}$ appear preliminary and require further scrutiny.

INTRODUCTION

Acute aortic syndromes (AASs) are deadly cardiovascular emergencies involving the thoracic aorta. They include acute aortic dissection, intramural aortic hematoma, penetrating aortic ulcer and aortic rupture.[1] AASs represent unique diagnostic challenges because they are relatively rare diseases (4-6 cases/100.000 individuals/year), but their presenting symptoms are unspecific and frequent in Emergency Department (ED) visits. For instance, chest pain accounts for ~6% of ED visits (8-10 million visits/year in the US), abdominal pain for ~6% and syncope for ~2%.[2-5] Conclusive diagnosis requires advanced imaging techniques, mostly contrast-enhanced computed tomography angiography (CTA), but owing to radiation, contrast exposure and resource limitations, CTA cannot be performed in all patients with AAS-compatible symptoms.[6] Consequently, decision on advanced imaging for suspected AASs is cumbersome, as shown by substantial variability in CTA ordering within emergency physicians, high misdiagnosis rates (up to 39%) and low diagnostic efficiency (as low as 2% of CTA exams turning out positive in North American series).[7-9]

For standardized clinical probability assessment of AASs, the reference tool indicated by guidelines is the aortic dissection detection risk score (ADD-RS), based on 12 risk factors organized in 3 categories (*supplementary table 1*).[10,11] Using the ADD-RS, patients can be classified in 3 risk-categories (ADD-RS=0 or low risk, ADD-RS=1 or intermediate risk, ADD-RS>1 or high risk), or in 2 risk-categories (ADD-RS≤1 or low probability, ADD-RS>1 or high-probability). In guidelines by the European Society of Cardiology (ESC) and American College of Cardiology (ACC)/American Heart Association (AHA), standardized probability assessment, in association with thorough physical examination, history collection and clinical reasoning, designs a pipeline for standardized diagnostic evaluation of stable patients with suspected AASs. However, the ADD-RS does not substitute clinical reasoning and is not recommended by the American College of Emergency Physicians (ACEP) in isolation.[12]

D-dimer (DD), a fibrinogen degradation product well-established for the rule-out of pulmonary embolism (PE), is also a highly sensitive and moderately specific biomarker of AASs.[13,14] The standard DD cutoff for AASs is 500 ng/mL. A key determinant of DD specificity is age, with a higher incidence of false positive results in elderly patients. For PE rule-out, application of an age-adjusted DD ($DD_{age\text{-}adj}$) increases specificity and efficiency without affecting sensitivity.[15] Two studies have reported that also for AASs, $DD_{age\text{-}adj}$ may increase specificity with a small trade-off in sensitivity.[16,17] A single cutoff for PE and AASs could be

very practical, as both conditions are invariably considered in differential diagnosis in patients with truncal pain and both imply decision on CTA.[18]

The rationale of integrating ADD-RS with DD testing is that very few cases of AASs are predicted to occur in patients with ADD-RS=0 or ≤1 and a negative DD test result.[14,19] In the present study, we aimed to provide a systematic review of studies evaluating the integration of ADD-RS with DD. For diagnostic variables with low statistical heterogeneity across studies, we aimed to determine pooled estimates. In order to also evaluate diagnostic bundles applying $DD_{age\text{-}adj}$, we obtained primary data from the investigators of the selected studies.

METHODS

Registration

The study protocol was registered in the international prospective register of systematic reviews (PROSPERO) website, with CRD42019137508. This study followed PRISMA statement and the recommendations included in the Cochrane Handbook Accuracy and in the PRISMA-DTA statement.[20,21] Institutional review board ethical approval was not needed because of the reviewing nature of this study.

Search strategy

In June 2019, we conducted a thorough online search on MEDLINE, EMBASE and Web Of Science databases. Detailed search strategies are presented in *table 1*. We subsequently hand-searched the reference lists of all articles identified in our searches and of systematic reviews and meta-analyses on this topic.

Inclusion/exclusion criteria

Two investigators independently reviewed the titles and abstracts of the studies to assess eligibility. The full text article of the potentially eligible articles was next obtained to evaluate inclusion/exclusion criteria. Any disagreement was solved by consensus. The study design was gathered from Asha *et al*., representing the reference meta-analysis for DD in AASs.[14] Studies were included if: (1) they were original research primarily assessing integration of ADD-RS with DD for the diagnosis of AASs; (2) they were cross-sectional diagnostic

studies; (3) prospective or retrospective enrollment was based on one or more AAS-compatible symptoms amongst chest pain, abdominal pain, back pain, syncope, perfusion deficit; (4) the ADD-RS was calculated; (5) the DD level was measured; (6) the diagnosis was confirmed or excluded with satisfactory criteria (advanced imaging with CTA, transesophageal echocardiography, magnetic resonance angiography, aortography, surgery or autopsy; in alternative, clinical case adjudication based on clinical data review and/or follow-up data); (7) absolute numbers of true positive, true negative, false positive, and false negative were reported or could be derived.

Studies were excluded if the design was case-control or case series due to high potential biases and the impossibility to calculate pre-test probability.[22] Conference abstracts were excluded because they are not peer-reviewed, the results may not be final and insufficient detail is provided for quality assessment.

Data extraction and analysis

The reporting of this systematic review follows the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA checklist, provided as *supplementary tables 2*).[21] Two reviewers extracted data independently from the selected articles. The extracted data included: first author, date of publication, study period, number of study sites, study setting, study design, inclusion/exclusion criteria, number of participants analyzed and excluded, D-dimer assay used, D-dimer reference range, reference standard used. They also extracted the study population characteristics (age, gender, time from symptom onset to evaluation), ADD-RS distribution, D-dimer level, AAS subtype, alternative final diagnoses made for patients without AASs, reporting the absolute number of true positive (TP), true negative (TN), false positive (FP), false negative (FN).

Two investigators independently assessed the quality of study design using the QUADAS-2 tool and the quality of reporting using the STARD tool.[23,24] QUADAS-2 assessment was done in compliance with the original background document.[23] For the domain "patient selection", we identified a high risk of bias if the sample was not consecutive, if the study wasn't done in the ED, if symptoms leading to patient inclusion did not include at least chest pain (representing the most common presenting symptom of AASs) and if patient enrollment was based on results of D-dimer or advanced imaging and not on clinical presentation. For the domain "index test", we identified a high risk of bias if the threshold of the index test wasn't prespecified or

if the result of the index test was interpreted after applying the reference standard. For the domain "reference standard", a high risk of bias was identified if patients were not subjected to advanced aortic imaging (CTA, transesophageal echocardiography, magnetic resonance angiography or aortography), surgery or autopsy. For patients not subjected to advanced imaging, surgery or autopsy, case adjudication based on independent clinical data review and/or follow-up data was considered satisfactory. For the domain "flow and timing", a high risk of bias was identified if studies included a significant (>5%) proportion of patients evaluated >14 days after symptom onset. Agreement between the reviewers was assessed with Cohen's k statistic. Types of diagnostic bias and anticipated skews in observed sensitivity/specificity were evaluated according to Kohn *et al*.[22]

Based on clinical reasoning and previous evidence, we planned to analyze the DD test results based on two different cutoffs: 500 ng/mL and an age-adjusted cutoff (DD_{adj}).[18] For the latter, the DD result was interpreted as follows: in patients younger than 50 years, an AAS was excluded in those with a DD value lower than 500 ng/mL. In patients aged 50 years or older, the DD test result was considered negative in those with a DD value lower than their age multiplied by 10. Briefly, DD_{adj} (ng/mL) was calculated as: age (years) × 10 ng/mL (with a minimum of 500 ng/mL).[15] To conduct an individual patient-level meta-analysis, the authors of all the selected studies were contacted to obtain missing data. For each study, a database was obtained reporting for each included patient, the age in years, the ADD-RS, the absolute DD level and the final diagnosis.

In the meta-analysis, we analyzed the performance of the following integrated strategies for diagnostic rule-out of AASs (*i.e.* if string satisfied, rule-out AASs): (1) ADD-RS≤1 and DD<500 ng/mL; (2) ADD-RS≤1 and $DD<DD_{age\text{-}adj}$; (3) ADD-RS=0 and DD<500 ng/mL; (4) ADD-RS=0 and $DD<DD_{age\text{-}adj}$. We built 2 x 2 contingency tables for each diagnostic strategy using the number of true positive (TP), false positive (FP), false negative (FN), true negative (TN). For negative likelihood ratio (LR) values of strategies with a sensitivity of 100%, contingency tables with zero value were handled by adding a 0.5 continuity correction and the 95% CI was estimated using a bootstrapping approach.[25] The failure rate was calculated as FN / (FN+TN), *i.e.* number of patients with AASs satisfying rule-out criteria divided by the total number of patients satisfying rule-out criteria.[26] The rule-out efficiency was calculated as (TN+FN) / (TP+FP+TN+FN), *i.e.* number of patients ruled-out by each integrated strategy divided by total number of patients tested. Heterogeneity was determined using

the Higgins' I^2. For variables showing non-significant heterogeneity, we calculated pooled values using fixed or random models as appropriate, based on inter- and intra-study variability.

The Pauker and Kassirer decision threshold model was applied to calculate two theoretical thresholds: a testing threshold (*i.e.* the probability of AAS at which there is no difference between performing the test and withholding the treatment) and a test-treatment threshold (*i.e.* the probability of AAS at which there is no difference between performing the test and administering the treatment).[27]

Statistical analysis was carried-out using Stata 13.1 (Stata Corp, College Station, Texas).

RESULTS

Literature search

Within 680 studies identified by the systematic database research, 12 studies were selected for full-text review (*figure 1*) and 4 studies met all the inclusion criteria.[16,28-30] Three studies were designed to investigate the diagnostic test characteristics of ADD-RS plus DD<500 ng/mL,[28-30] and one investigated the integration of ADD-RS with DD<$DD_{age-adj}$.[16] The study characteristics are summarized in *table 2* and the final diagnoses of the participant patients are detailed in *supplementary table 3*. The case-mix of AASs was similar amongst studies, with acute aortic dissection representing the most frequent subtype and intramural aortic hematoma or penetrating aortic ulcer accounting for most of the other cases. Some specificities were found in the study by Kotani *et al*., which included a larger number of patients with complicated aneurysms (ruptured/with impending-rupture or infectious).[16] This study also reflects the higher prevalence of intramural hematomas in Japan and Asia. The pooled prevalence of AASs (mean ± SD) across the four studies was 18.0% ± 5.3%, which is substantially higher than reported in North American ED series, but also substantially lower than in most diagnostic biomarker studies, including the first key prospective multicenter study of D-dimer.[9,31-33] The pooled prevalence of "classic" acute aortic dissection was 67.3%, of intramural aortic hematoma was 18.7% and of penetrating aortic ulcer was 6.8%. A higher observed prevalence of AASs than in general ED practice could lead to spectrum bias (falsely raising sensitivity).

The study by Nazerian *et al*. (2018) was the only prospective multicenter study. Its primary aim was to define the failure rate of a diagnostic rule-out strategy integrating ADD-RS (=0 or ≤1) with DD<500

ng/mL.[30] A secondary analysis applying $DD_{age-adj}$ has also been published.[17] In this trial, the gold standards for case-adjudication were conclusive aortic imaging (by CTA, transesophageal echocardiography or magnetic resonance angiography), surgery or autopsy. Enrollment preceded final decision on aortic imaging, and patients who were not subjected to any of these gold standards during the ED visit were subjected to 14-day follow-up. Patients or family members were interviewed by telephone with a structured questionnaire or underwent an outpatient visit after ED discharge, evaluating diagnosis of any aortic disease, subsequent ED visits, hospital admission and death. Reviewers defining case adjudication had access to hospital charts and discharge documents. Gold standard imaging was obtained during the index visit in 45% of patients, and during follow-up in 1.6% of patients. Patients dismissed from the ED and with a negative follow-up were 34.3%, potentially leading to differential verification bias, with a decrease in observed sensitivity and specificity. During follow-up, 2 patients were lost and 3 cases of AASs were diagnosed. Three patients died without advanced imaging or surgery. They all had a positive DD test result and therefore could not be regarded as potential FN cases.

The studies by Nazerian *et al*. (2014), Gorla *et al*. and Kotani *et al*. were retrospective.[16,28,29] In these studies, an exact time definition of symptoms triggering enrollment was not reported by the authors. This raises concern about the potential inclusion of patients with non-acute symptoms, in whom the chance of FN cases is higher. In the study by Nazerian *et al*. (2014), data were obtained from a registry of ED patients undergoing advanced aortic imaging for clinically suspected AAS.[28] The study cohort largely overlapped with another study from the same groups which focused on validation of the ADD-RS *per se*.[34] For 29% of enrolled participants, a DD test result was not available, leading to patient exclusion. This could introduce partial verification bias (which could raise sensitivity), but the characteristics of the patients in the included and excluded groups were similar. A *post*-hoc analysis showed that only 17 patients (1.6%) presented with history of pain >14 days; 4 of them had an AAS. One of these patients (symptoms for 15 days), had a normal D-dimer. However, exact time data was missing for 39.2% of the enrolled patients.

In Gorla *et al*.,[29] patients were enrolled if they were admitted to the ED for chest pain and if they were subjected to a DD assay. These criteria could bias against atypical presentations not involving chest pain. The clinical judgment of the physician ordering DD was not recorded. Hence, PE and not AAS could have represented the chief differential diagnosis in some patients. Indeed, the rate of PE in this cohort was 14.5%,

significantly higher than in the other studies. The authors declare that in study patients, CTA was used *per* guidelines, based on clinical judgment and on DD test result. All diagnoses of AASs were confirmed by advanced aortic imaging. Since the actual number of patients subjected to advanced imaging is unknown and clinical case adjudication was not based on a pre-specified follow-up, observed sensitivity could be raised due to differential verification bias.

In the study by Kotani *et al*.,[16] patients were also enrolled if they presented with acute chest pain and if they received a DD assay. The exact time interval from symptom onset to sampling was not presented, and the DD assay was used *per* a pre-specified hospital protocol not detailed in the manuscript. The analysis was conducted only on patients admitted to hospital after the ED visit, while patients dismissed from the ED were excluded. This could lead to spectrum bias, raising sensitivity in the enrolled sample. Restriction to admitted patients potentially biases towards a more clinically severe population, while rule-out strategies ideally apply to patients in whom early ED discharge represents a meaningful option. However, the final prevalence of AASs was 13.9%, indicating adequate representation of low-probability patients. Additional exclusion criteria were ST elevation on ECG and hemodynamic instability. Both criteria are in line with ESC recommendations, as patients with these clinical characteristics are not amenable to rule-out criteria.[11] The DD assay was interpreted using the $DD_{age\text{-}adj}$ cutoff. As in Gorla *et al*., enrollment criteria focused on chest pain, excluding alternative clinical presentations and likely included patients with a clinical suspicion of PE and not only of AASs. However, the prevalence of PE was generally low (3.8%), while the prevalence of acute coronary syndromes was the highest, indicating potential bias towards coronary artery disease.

Quality assessment

The quality assessment (QA) conducted using the QUADAS-2 is shown in *table 3* and in *supplementary figure 1*. For only one study, the judgment was "low" in all 7 domains, indicating an overall low risk of bias and concern regarding applicability.[30] In one study, the judgment was "low" in 3 of 7 domains.[29]

The quality of reporting of the included studies, analyzed according to the STARD 2015 statement, is detailed in *supplementary table 4*. Most studies showed suboptimal quality regarding type of sample enrollment, how missing data on the index test and reference standard were handled, sample size calculation,

whether any clinical intervention was done between the index test and the reference standard, study registration and accessibility of the full protocol. The agreement between the reviewers for components of the study quality assessment tools was good (κ=0.67, 95% CI 0.54 – 0.80).[35]

Meta-analysis

A total of 3804 patients were included in the meta-analysis, including 675 (17.7%) with AASs. To evaluate strategies integrating either the 500 ng/mL or the $DD_{age-adj}$ cutoff, individual patient-level data were used. Contingency tables and coupled forest plots were obtained (*figure 2*). For all strategies, statistical heterogeneity was negligible for sensitivity (I^2=0%) and significant for specificity values. Subanalyses excluding patients with ADD-RS=0, shown in *supplementary tables 5-6*, indicated that results were not substantially affected by inclusion of patients at lowest pre-test probability of AASs. Negative and positive likelihood ratio (LR) values of the diagnostic strategies are shown in *figure 3*. Heterogeneity was negligible for the negative LR (I^2=0%) and significant for the positive LR values of the diagnostic strategies.

Forest plots of failure rate and efficiency values are shown in *figures 4-5*. Failure rate values had low to moderate heterogeneity for ADD-RS=0 and DD<500 ng/mL (I^2=38.1%), ADD-RS=0 and DD<$DD_{age-adj}$ (I^2=28%), ADD-RS≤1 and DD<500 ng/mL (I^2=39%); heterogeneity was significant for ADD-RS≤1 plus DD<$DD_{age-adj}$ (I^2=84.4%). Efficiency values had significant heterogeneity for all diagnostic strategies.

Pooled estimates of diagnostic variables underlying diagnostic rule-out (sensitivity, negative LR and failure rate) and showing non-significant heterogeneity across studies are summarized in *table 4*. Diagnostic variables showing high heterogeneity were not to reported, as limited inference on pooled values can be done. For ADD-RS=0 and DD<500 ng/mL, pooled sensitivity was 99.9% (99.3–100%), negative LR 0.032 (0–0.086) and failure rate 0.1% (0–0.3%). For ADD=0 and DD<$DD_{age-adj}$ sensitivity was 99.9% (99.3–100%), negative LR 0.027 (0–0.081) and failure rate 0.1% (0–0.2%). For ADD-RS≤1 and DD<500 ng/mL, sensitivity was 98.9% (97.9–99.9%), negative LR 0.025 (0.001–0.049) and failure rate 0.6% (0.2–0.9%). For ADD-RS≤1 and DD<$DD_{age-adj}$, sensitivity was 97.6% (96.3–98.9%) and negative LR 0.048 (0.022–0.074). For this strategy, pooled failure was not computed due to significant heterogeneity.

Test-treatment threshold

Test-treatment thresholds were calculated for diagnostic strategies including DD<500 ng/ml (*supplementary figure 2*). According to this model, the ADD-RS=0 and DD<500 ng/mL strategy should be performed if the clinical probability of AASs is between 1.7% and 23.2%, while the ADD-RS≤1 and DD<500 ng/mL strategy should be performed when the pretest probability is between 1.1% and 44.8%.

DISCUSSION

We provide a systematic review and summary of studies assessing integration of ADD-RS with DD for diagnosis of AASs. Only four papers satisfied the pre-defined inclusion criteria, underlying the relative paucity of data. However, the total number of included patients was substantial (n=3804). All studies post-dated the latest guidelines of the American Heart Association and the European Society of Cardiology, and only one (Nazerian *et al*. 2014) was cited in the latest clinical policy of the American College of Emergency Physicians.[11,12,36] One was a prospective multicenter trial, while the other 3 were retrospective studies. All were performed in the ED and mostly involved patients with chest pain, but inclusion criteria partly differed. This key limit reflects the absence of a standard definition of patients suspected of having AASs and amenable to rule-out strategies. Therefore, methodological and clinical heterogeneity between available studies mandate caution in efforts to pool and summarize data.

Significant statistical heterogeneity was found for specificity, positive LR and efficiency. This likely reflects differences in the clinical case-mix of study cohorts. For these variables, data pooling could be misleading and was therefore omitted. AASs cases were instead homogenous across studies, thus leading to negligible statistical heterogeneity for sensitivity and negative LR values and allowing meaningful data pooling for these variables. Also in a previous meta-analysis of high-quality studies (which also included Nazerian *et al*. 2014), the heterogeneity was low for sensitivity and negative LR, and substantial for specificity and positive LR.[14]

Acquisition of primary data allowed us to evaluate diagnostic strategies incorporating also $DD_{age\text{-}adj}$, already in use for PE rule-out. When using ADD-RS=0, $DD_{age\text{-}adj}$ provided pooled sensitivity and negative LR values similar to those of the "classical" 500 ng/mL cutoff. Instead, when using ADD-RS≤1, DD<500 ng/mL outperformed DD<$DD_{age\text{-}adj}$ in terms of pooled sensitivity and negative LR. This data suggests that DD<$DD_{age\text{-}}$

$_{adj}$ could be evaluated in further studies only if the pre-test probability is presumed to be very low. $DD_{age\text{-}adj}$ might provide increased specificity over 500 ng/mL, but the statistical heterogeneity found across studies does not allow any conclusion.

Consensus is lacking on what should reproducibly define a clinical suspicion of AASs. Hence, differences between physicians and centers can be profound. In North American retrospective series of patients undergoing CTA for suspected AAS, the prevalence of AASs was ~3%.[8,37] In a vast out-of-hospital study evaluating the ADD-RS in non-traumatic emergencies, the prevalence of AASs was 0.9%.[38] In the studies reviewed herein, the prevalence of AASs was 13 to 23%. Application of rule-out strategies to patient populations at lower pre-test probability of AASs is expected to result in lower failure rates, with a trade-off in efficiency.

Caution is needed when considering application of ADD-RS and D-dimer based strategies in clinical practice. First, ADD-RS, a decision rule derived from a retrospective register of AASs, has low specificity.[37] In addition, ADD-RS derivation methods have not been published, and it is currently unknown whether use of the ADD-RS provides any advantage in terms of diagnostic accuracy and of CTA ordering, as compared to clinical gestalt.[39,40] In the future, focused ED-centered studies may provide alternative and more specific probability assessment tools. Second, D-dimer also lacks specificity. Therefore, indiscriminate application of ADD-RS and D-dimer to unselected ED patients with AAS-compatible symptom(s) would paradoxically increase the number of CTA ordered. Such slippery slope must be avoided.[41]

Based on previous data, in terms of specificity, we speculate that the ADD-RS/D-dimer rule-out pathway could best apply to stable patients with ADD-RS=1 owing to clinical manifestations providing *per se* higher specificity (*i.e.* pulse deficit, neurological deficit, aortic valve insufficiency).[42] Caution is needed in patients with hypotension, which also potentially defines clinical instability and might prompt towards a fast-track for advanced imaging irrespective of D-dimer test results. However, in clinical practice, most cases with ADD-RS=1 will be driven by pain features (severe, sudden, ripping pain), providing higher sensitivity but lower specificity. To maximize benefits, a pragmatic approach could be to request D-dimer only after three-dimensional evaluation of clinical history, physical examination, first-line imaging and blood test results, in patients still lacking a clear alternative diagnosis, or in whom rule-out of AASs is considered imperative for decision on hospital admission versus discharge or administration of anticoagulant/antiplatelet therapies,

which could be harmful in presence of an AAS.

LIMITATIONS

Only one study (49% of patients) was judged to provide a low risk of bias/applicability concerns.[30] Two studies (42% of patients)[16,28] had issues in one of the QUADAS domains, and one study (10% of patients) had a generally lower quality profile.[29] In one study, the case-mix of AASs slightly differed, with fewer cases of acute aortic dissections and higher prevalence of the other forms.[16] Overall, the potential bias types most frequently encountered were: (1) partial verification bias, due to patients excluded because discharged from the ED or due to unavailable DD test result (leading to potential upward skew in sensitivity and downward skew in specificity), and (2) differential verification bias, due to inclusion of patients subjected to clinical follow-up without advanced aortic imaging (leading to potential downward skew in sensitivity and specificity). The accuracy of DD for diagnosis of AASs may also slightly differ amongst subtypes, with higher risk of false negative cases in patients with intramural hematomas and focal dissections.[43,44] Therefore, methodological and clinical heterogeneity between available studies mandate caution in data pooling and summarization.

A key issue affecting two studies (24% of patients) is that the authors selected patients with chest pain and a DD test result, potentially also introducing individuals with suspected PE.[16,29] In clinical terms, this aspect may be secondary, because both PE and AASs are typically considered in differential diagnosis, share DD as the key biomarker and require CTA for conclusive diagnosis. A suspicion of AAS by the attending physicians was clearly defined in two studies (76% of patients) led by the same primary investigators.[28,30] This might limit external validity.

With the exception of the ADvISED trial, there was general uncertainty about the timing of the index test. Hence, a minority of patients with symptoms dating >14 days were possibly enrolled, including few cases of AASs in their subacute or chronic phase. Since D-dimer levels tend to decrease over time after development of AASs, this is expected to increase the number of patients with AASs presenting as FN (differential verification bias, with potential downward skew in estimates of sensitivity and failure rate).[45]

CONCLUSIONS

Only 4 studies have evaluated integration of ADD-RS with DD for diagnosis of AASs, with methodological differences that must be carefully considered. However, the total number of included patients is reasonably large (n=3804), and negligible heterogeneity was found for sensitivity and negative LR values. Available studies consistently show that ADD-RS=0 or ≤1 plus DD<500 ng/mL are highly sensitive diagnostic strategies and support their reliability for rule-out of AASs. For $DD_{age\text{-}adj}$, available data appear largely preliminary and further studies are required. Nonetheless, further prospective trials, especially in low prevalence populations, are needed to confirm the results of this meta-analysis.

REFERENCES

1. Bossone E, LaBounty TM, Eagle KA. Acute aortic syndromes: diagnosis and management, an update. Eur Heart J 2018;39:739-49d.

2. Goodacre S, Cross E, Arnold J, Angelini K, Capewell S, Nicholl J. The health care burden of acute chest pain. Heart 2005;91:229-30.

3. Alshekhlee A, Shen WK, Mackall J, Chelimsky TC. Incidence and mortality rates of syncope in the United States. Am J Med 2009;122:181-8.

4. Amsterdam EA, Kirk JD, Bluemke DA, et al. Testing of low-risk patients presenting to the emergency department with chest pain: a scientific statement from the American Heart Association. Circulation 2010;122:1756-76.

5. Owens PL, Barrett ML, Gibson TB, Andrews RM, Weinick RM, Mutter RL. Emergency department care in the United States: a profile of national data sources. Ann Emerg Med 2010;56:150-65.

6. Klompas M. Does this patient have an acute thoracic aortic dissection? JAMA 2002;287:2262-72.

7. Hansen MS, Nogareda GJ, Hutchison SJ. Frequency of and inappropriate treatment of misdiagnosis of acute aortic dissection. Am J Cardiol 2007;99:852-6.

8. Lovy AJ, Bellin E, Levsky JM, Esses D, Haramati LB. Preliminary development of a clinical decision rule for acute aortic syndromes. Am J Emerg Med 2013;31:1546-50.

9. Ohle R, Anjum O, Bleeker H, Wells G, Perry JJ. Variation in emergency department use of computed tomography for investigation of acute aortic dissection. Emerg Radiol 2018;25:293-8.

10. Rogers AM, Hermann LK, Booher AM, et al. Sensitivity of the aortic dissection detection risk score, a novel guideline-based tool for identification of acute aortic dissection at initial presentation: results from the international registry of acute aortic dissection. Circulation 2011;123:2213-8.

11. Erbel R, Aboyans V, Boileau C, et al. 2014 ESC Guidelines on the diagnosis and treatment of aortic diseases: Document covering acute and chronic aortic diseases of the thoracic and abdominal aorta of the adult. The Task Force for the Diagnosis and Treatment of Aortic Diseases of the European Society of Cardiology (ESC). Eur Heart J 2014;35:2873-926.

12. Diercks DB, Promes SB, Schuur JD, Shah K, Valente JH, Cantrill SV. Clinical policy: critical issues in the evaluation and management of adult patients with suspected acute nontraumatic thoracic aortic dissection. Ann Emerg Med 2015;65:32-42 e12.

13. Konstantinides SV, Torbicki A, Agnelli G, et al. 2014 ESC guidelines on the diagnosis and management of acute pulmonary embolism. Eur Heart J 2014;35:3033-69, 69a-69k.

14. Asha SE, Miers JW. A Systematic Review and Meta-analysis of D-dimer as a Rule-out Test for Suspected Acute Aortic Dissection. Ann Emerg Med 2015;66:368-78.

15. Righini M, Van Es J, Den Exter PL, et al. Age-adjusted D-dimer cutoff levels to rule out pulmonary embolism: the ADJUST-PE study. JAMA 2014;311:1117-24.

16. Kotani Y, Toyofuku M, Tamura T, et al. Validation of the diagnostic utility of D-dimer measurement in patients with acute aortic syndrome. Eur Heart J Acute Cardiovasc Care 2017;6:223-31.

17. Morello F, Mueller C, Soeiro AM, et al. Response by Morello et al to Letters Regarding Article, "Diagnostic Accuracy of the Aortic Dissection Detection Risk Score Plus D-Dimer for Acute Aortic Syndromes: The ADvISED Prospective Multicenter Study". Circulation 2018;138:448-9.

18. Suzuki T, Eagle KA. Biomarker-Assisted Diagnosis of Acute Aortic Dissection. Circulation 2018;137:270-2.

19. Baez AA, Cochon L. Improved rule-out diagnostic gain with a combined aortic dissection detection risk score and D-dimer Bayesian decision support scheme. J Crit Care 2017;37:56-9.

20. Deeks JJ, Bossuyt PM, Gatsonis C. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. 2013.

21. McInnes MDF, Moher D, Thombs BD, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA 2018;319:388-96.

22. Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. Acad Emerg Med 2013;20:1194-206.

23. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.

24. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.

25. Marill KA, Chang Y, Wong KF, Friedman AB. Estimating negative likelihood ratio confidence when test sensitivity is 100%: A bootstrapping approach. Stat Methods Med Res 2017;26:1936-48.

26. van Es N, van der Hulle T, van Es J, et al. Wells Rule and d-Dimer Testing to Rule Out Pulmonary Embolism: A Systematic Review and Individual-Patient Data Meta-analysis. Ann Intern Med 2016;165:253-61.

27. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. N Engl J Med 1980;302:1109-17.

28. Nazerian P, Morello F, Vanni S, et al. Combined use of aortic dissection detection risk score and D-dimer in the diagnostic workup of suspected acute aortic dissection. Int J Cardiol 2014;175:78-82.

29. Gorla R, Erbel R, Kahlert P, et al. Accuracy of a diagnostic strategy combining aortic dissection detection risk score and D-dimer levels in patients with suspected acute aortic syndrome. Eur Heart J Acute Cardiovasc Care 2017;6:371-8.

30. Nazerian P, Mueller C, Soeiro AM, et al. Diagnostic Accuracy of the Aortic Dissection Detection Risk Score Plus D-Dimer for Acute Aortic Syndromes: The ADvISED Prospective Multicenter Study. Circulation 2018;137:250-8.

31. Suzuki T, Distante A, Zizza A, et al. Diagnosis of acute aortic dissection by D-dimer: the International Registry of Acute Aortic Dissection Substudy on Biomarkers (IRAD-Bio) experience. Circulation 2009;119:2702-7.

32. Taylor RA, Iyer NS. A decision analysis to determine a testing threshold for computed tomographic angiography and d-dimer in the evaluation of aortic dissection. Am J Emerg Med 2013;31:1047-55.

33. Wang Y, Tan X, Gao H, et al. Magnitude of Soluble ST2 as a Novel Biomarker for Acute Aortic Dissection. Circulation 2018;137:259-69.

34. Nazerian P, Giachino F, Vanni S, et al. Diagnostic performance of the aortic dissection detection risk score in patients with suspected acute aortic dissection. Eur Heart J Acute Cardiovasc Care 2014;3:373-81.

35. Byrt T. How good is that agreement? Epidemiology 1996;7:561.

36. Hiratzka LF, Bakris GL, Beckman JA, et al. 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM guidelines for the diagnosis and management of patients with Thoracic Aortic Disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, American College of Radiology, American Stroke Association, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society of Interventional Radiology, Society of Thoracic Surgeons, and Society for Vascular Medicine. Circulation 2010;121:e266-369.

37. Ohle R, Anjum O, Bleeker H, McIsaac S. What Is the Specificity of the Aortic Dissection Detection Risk Score in a Low-prevalence Population? Acad Emerg Med 2019;26:632-8.

38. Yamashita A, Maeda T, Kita Y, et al. The impact of prehospital assessment and EMS transport of acute aortic syndrome patients. Am J Emerg Med 2018;36:1188-94.

39. Hill JM, Murphy TG, Fermann GJ. Aortic Dissection Detection Risk Score: A Clinical Decision Rule That Needs Some Parenting. Acad Emerg Med 2019;26:695-7.

40. Smith LM, Miller CD. Acute Aortic Dissection: Is There Something Better than Physician Gestalt? Acad Emerg Med 2018;25:464-6.

41. Fatovich DM. The inverted U curve and emergency medicine: Overdiagnosis and the law of unintended consequences. Emerg Med Australas 2016;28:480-2.

42. Ohle R, Kareemi HK, Wells G, Perry JJ. Clinical Examination for Acute Aortic Dissection: A Systematic Review and Meta-analysis. Acad Emerg Med 2018;25:397-412.

43. Ohlmann P, Faure A, Morel O, et al. Lower circulating Sta-Liatest D-Di levels in patients with aortic intramural hematoma compared with classical aortic dissection. Crit Care Med 2009;37:899-901.

44. Gorla R, Erbel R, Kahlert P, et al. Diagnostic role and prognostic implications of D-dimer in different classes of acute aortic syndromes. Eur Heart J Acute Cardiovasc Care 2017;6:379-88.

45. Eggebrecht H, Naber CK, Bruch C, et al. Value of plasma fibrin D-dimers for detection of acute aortic dissection. J Am Coll Cardiol 2004;44:804-9.

FIGURE LEGENDS

Figure 1. PRISMA flow diagram of study search and selection.

Figure 2. Contingency tables and coupled forest plots of sensitivity and specificity values. Heterogeneity was determined using the Higgins' I^2.

Figure 3. Forest plots of the negative and positive likelihood ratio values. Heterogeneity was determined using the Higgins' I^2.

Figure 4. Forest plots of the failure rate values. Heterogeneity was determined using the Higgins' I^2.

Figure 5. Forest plots of the efficiency values. Heterogeneity was determined using the Higgins' I^2.

TABLES

Table 1. Detailed database search strategies.

Literature database	Search query
MEDLINE	((((("Aneurysm, Dissecting"[Mesh]) AND "Fibrin Fibrinogen Degradation Products"[Mesh]) OR (acute aortic syndrome AND D-dimer))) OR "Aortic Dissection Detection Risk Score")
EMBASE	(('acute aortic syndrome'/exp OR 'acute aortic syndrome' OR 'aortic dissection'/exp OR 'aortic dissection') AND ('d dimer'/exp OR 'd dimer') OR 'aortic dissection detection risk score') NOT 'conference abstract':it NOT review:it NOT letter:it
Web Of Science	TOPIC: (("acute aortic syndrome" OR "aortic dissection" OR "dissecting aneurysm" OR "Aortic Dissection Detection Risk Score") AND ("D-dimer" OR "Fibrin Degradation Product"))

Table 2. Characteristics of included studies.

	Nazerian, 2014	**Gorla, 2017**	**Kotani, 2017**	**Nazerian, 2018**
Study period	01/2008 to 03/2013	01/2001 to 05/2014	01/2011 to 04/2014	01/2014 to 12/2016
N. of study sites	2	1	1	6
Setting country hospital department	 IT Large referral ED	 GE NR ED	 JA Large referral ED	 BR, SW, GE, IT Large referral ED
Participants, N (% of enrolled)	1035 (71%)	376 (100%)	545 (61.4%)	1848 (99.9%)
Participants excluded for unavailable index test, N (%)	420 (29%)	0 (0%)	66 (6.9%)	48 (2.5%)
AASs, N (% enroll.) AD, N (% AAS) IMH, N (% AAS) PAU, N (% AAS) other, N (% AAS)	233 (22.5%) 199 (85.4%) 31 (13.3%) 3 (1.3%) 0	85 (22.6%) 61 (71.8%) 11 (12.9%) 13 (15.3%) 0	123 (13.9%) 47 (38.2%) 42 (34.1%) 8 (6.5%) 26 (21.1%)[#]	241 (13%) 178 (73.9%) 35 (14.5%) 10 (4.1%) 18 (7.5%)[¶]
ADD-RS, N (N, % with AAS) 0 1 2-3	 322 (19, 5.9%) 508 (133, 26.2%) 205 (81, 39.5%)	 189 (1, 0.5%) 130 (30, 23.1%) 57 (54, 94.7%)	 75 (4, 5.3%)[§] 399 (88, 22.1%)[§] 71 (24, 33.8%)[§]	 437 (12, 2.7%) 1070 (96, 9.0%) 341 (133, 39.0%)
Study design	Prospective enrollment, retrospective analysis	Retrospective	Retrospective	Prospective
Inclusion criteria	chest/back/ abdominal pain, syncope or perfusion deficit + alt-D not established + clinical suspicion leading to CTA	chest pain + D-dimer available at presentation	acute chest pain + admission to hospital + D-dimer available	chest/back/abdominal pain, syncope or perfusion deficit + clinical suspicion
Exclusion criteria	NR	NR	hemodynamic instability, STEMI, ED discharge, death	primary trauma, unwillingness or inadequacy to participate

			in ED, referral to other hospital	
Patient sampling	NR	NR	NR	Consecutive
Reference standard	CTA	unspecified advanced imaging study	CTA	CTA, TEE, MRA, surgery or autopsy; if unavailable, 14-day clinical follow-up
Age, y mean (SD)	67 (14%)	63 (12%)	70 (14%)	62 (12%)
Male, %	66%	61%	63.4%	62.3%
Duration of symptoms, hours	48 (7-96)*§	NR	82% <24 h	7.5 (2-30)*
D-dimer assay	HemosIL D-Dimer HS, STA®-Liatest® D-Di	Innovance® D-Dimer	Liatest D-dimer, Hexamate D-dimer	HemosIL D-Dimer HS, STA®-Liatest® D-Di, TriniLIA D-Dimer, Innovance® D-Dimer
D-dimer cutoff (ng/ml)	<500 ng/ml	≤500 ng/ml	if age ≤50y: < 500 ng/ml if age >50y: < (age x 10) ng/ml	<500 ng/ml
D-dimer, test char. sensitivity specificity	 98.3% 35.9%	 97.6% 63.2%	 96.0% 58.0%	 96.7% 64.0%

AAS: acute aortic syndrome; AD: aortic dissection; alt-D: alternative diagnosis; BR: Brazil; CTA: computed tomography angiography; GE: Germany; IMH: intramural aortic hematoma; IQR: interquartile range; IT: Italy; JA: Japan; MRA: magnetic resonance angiography; NR: not reported; PAU: penetrating aortic ulcer; SD: standard deviation; SW: Switzerland; TEE: transesophageal echocardiography.

*Values reported as median (IQR). §Original data provided by the authors for the present analysis and not included in the original manuscript. #Includes: ruptured aortic aneurysm (7.3%), impending rupture of aortic aneurysm (10.6%), infectious aortic aneurysm (3.2%). ¶Includes only spontaneous (non-traumatic) rupture of thoracic aorta.

Table 3. Assessment of study quality according to QUADAS-2.[23]

Study	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Nazerian, 2014	L	L	L	U	L	L	L
Gorla, 2017	H	L	U	U	L	L	U
Kotani, 2017	H	L	L	L	L	L	L
Nazerian, 2018	L	L	L	L	L	L	L

L =Low-risk, H =High-risk, U =Unclear

Table 4. Pooled estimates of diagnostic variables underlying diagnostic rule-out.

	Sensitivity (%, 95% CI)	**Negative LR (95% CI)**	**Failure rate (%, 95% CI)**
ADD-RS=0 and DD<500 ng/mL	99.9% (99.3–100%)	0.032 (0-0.086)	0.1% (0–0.3%) 1 in 1000 (333–∞)
I-squared, *p*	0%, 0.95	0%, 0.64	38.1%, 0.18
ADD-RS=0 and DD<$DD_{age\text{-}adj}$	99.9% (99.3–100%)	0.027 (0-0.081)	0.1% (0–0.2%) 1 in 1000 (500–∞)
I-squared, *p*	0%, 0.95	0%, 0.77	28%, 0.24
ADD-RS≤1 and DD<500 ng/mL	98.9% (97.9–99.9%)	0.025 (0.001–0.049)	0.6% (0.2–0.9%) 1 in 167 (111–500)
I-squared, *p*	0%, 0.91	0%, 0.98	39%, 0.19
ADD-RS≤1 and DD<$DD_{age\text{-}adj}$	97.6% (96.3–98.9%)	0.048 (0.022–0.074)	n.a.
I-squared, *p*	0%, 0.86	0%, 0.56	84.4%, <0.001

LR: likelihood ratio; n.a.=not applicable due to significant heterogeneity.

ABSTRACT

Objectives. In patients at low clinical probability of acute aortic syndromes (AASs), decision on advanced aortic imaging is cumbersome. Integration of the aortic dissection detection risk score (ADD-RS) with D-dimer (DD) ~~designs a~~provides a potential pipeline for standardized diagnostic rule-out. We systematically reviewed and summarized supporting data.

Methods. Cross-sectional studies assessing integration of ADD-RS with DD for diagnosis of AASs were identified on MEDLINE, EMBASE and Web Of Science databases. Two reviewers independently screened articles, assessed quality and extracted data. The quality of design and reporting was evaluated with the QUADAS-2 and STARD tools. Individual patient data were obtained, to allow analysis of both conventional (500 ng/mL) and age-adjusted ($DD_{age\text{-}adj}$) DD cutoffs. Data were summarized for 4 diagnostic strategies combining ADD-RS=0 or ≤1, with DD<500 ng/mL or <$DD_{age\text{-}adj}$. The statistical heterogeneity of the diagnostic variables was estimated with Higgins' I^2. Pooled values were calculated for variables showing non-significant heterogeneity.

Results. After screening of 680 studies, 4 articles (including a total of 3804 patients) met inclusion criteria. One prospective study provided a low risk of bias/applicability concerns, while methodological limitations were found in the other 3 retrospective studies. Statistical heterogeneity was negligible for sensitivity and negative likelihood ratio (LR) values, and significant for specificity and positive LR values of all diagnostic strategies. Pooled sensitivity was 99.9% (95%CI 99.3–100%, I^2=0) for ADD-RS=0 and DD<500 ng/mL or <$DD_{age\text{-}adj}$, 98.9% (95%CI 97.9–99.9%, I^2=0) for ADD-RS≤1 and DD<500 ng/mL and 97.6% (95%CI 96.3–98.9%, I^2=0) for ADD-RS≤1 and DD<$DD_{age\text{-}adj}$.

Conclusions. Despite methodological limitations, integration of ADD-RS=0 or ≤1 with DD<500 ng/mL shows negligible heterogeneity and consistently high sensitivity across studies, thus supporting reliability for diagnostic rule-out of AASs. Data supporting ADD-RS=0 plus $DD_{age\text{-}adj}$ appear preliminary and require further scrutiny.

INTRODUCTION

Acute aortic syndromes (AASs) are deadly cardiovascular emergencies involving the thoracic aorta. They include acute aortic dissection, intramural aortic hematoma, penetrating aortic ulcer and aortic rupture.[1] AASs represent unique diagnostic challenges because they are relatively rare diseases (4-6 cases/100.000 individuals/year), but their presenting symptoms are unspecific and frequent in Emergency Department (ED) visits. For instance, chest pain accounts for ~6% of ED visits (8-10 million visits/year in the US), abdominal pain for ~6% and syncope for ~2%.[2-5] Conclusive diagnosis requires advanced imaging techniques, mostly contrast-enhanced computed tomography angiography (CTA), but owing to radiation, contrast exposure and resource limitations, CTA cannot be performed in all patients with AAS-compatible symptoms.[6] Consequently, decision on advanced imaging for suspected AASs is cumbersome, as shown by substantial variability in CTA ordering within emergency physicians, high misdiagnosis rates (up to 39%) and low diagnostic efficiency (as low as 2% of CTA exams turning out positive in North American series).[7-9]

For standardized clinical probability assessment of AASs, the reference tool indicated by guidelines is the aortic dissection detection risk score (ADD-RS), based on 12 risk factors organized in 3 categories (*supplementary table 1*).[10,11] Using the ADD-RS, patients can be classified in 3 risk-categories (ADD-RS=0 or low risk, ADD-RS=1 or intermediate risk, ADD-RS>1 or high risk), or in 2 risk-categories (ADD-RS≤1 or low probability, ADD-RS>1 or high-probability). In guidelines by the European Society of Cardiology (ESC) and American College of Cardiology (ACC)/American Heart Association (AHA), standardized probability assessment, in association with thorough physical examination, history collection and clinical reasoning, designs a pipeline for standardized diagnostic evaluation of stable patients with suspected AASs. However, the ADD-RS does not substitute clinical reasoning and is not recommended by the American College of Emergency Physicians (ACEP) in isolation.[12]

D-dimer (DD), a fibrinogen degradation product well-established for the rule-out of pulmonary embolism (PE), is also a highly sensitive and moderately specific biomarker of AASs.[13,14] The standard DD cutoff for AASs is 500 ng/mL. A key determinant of DD specificity is age, with a higher incidence of false positive results in elderly patients. For PE rule-out, application of an age-adjusted DD ($DD_{age\text{-}adj}$) increases specificity and efficiency without affecting sensitivity.[15] Two studies have reported that also for AASs, $DD_{age\text{-}adj}$ may increase specificity with a small trade-off in sensitivity.[16,17] A single cutoff for PE and AASs could be

very practical, as both conditions are invariably considered in differential diagnosis in patients with truncal pain and both imply decision on CTA.[18]

The rationale of integrating ADD-RS with DD testing is that very few cases of AASs are predicted to occur in patients with ADD-RS=0 or ≤1 and a negative DD test result.[14,19] In the present study, we aimed to provide a systematic review of studies evaluating the integration of ADD-RS with DD. For diagnostic variables with low statistical heterogeneity across studies, we aimed to determine pooled estimates. In order to also evaluate diagnostic bundles applying $DD_{age\text{-}adj}$, we obtained primary data from the investigators of the selected studies.

METHODS

Registration

The study protocol was registered in the international prospective register of systematic reviews (PROSPERO) website, with CRD42019137508. This study followed PRISMA statement and the recommendations included in the Cochrane Handbook Accuracy and in the PRISMA-DTA statement.[20,21] Institutional review board ethical approval was not needed because of the reviewing nature of this study.

Search strategy

In June 2019, we conducted a thorough online search on MEDLINE, EMBASE and Web Of Science databases. Detailed search strategies are presented in *table 1*. We subsequently hand-searched the reference lists of all articles identified in our searches and of systematic reviews and meta-analyses on this topic.

Inclusion/exclusion criteria

Two investigators independently reviewed the titles and abstracts of the studies to assess eligibility. The full text article of the potentially eligible articles was next obtained to evaluate inclusion/exclusion criteria. Any disagreement was solved by consensus. The study design was gathered from Asha *et al*., representing the reference meta-analysis for DD in AASs.[14] Studies were included if: (1) they were original research primarily assessing integration of ADD-RS with DD for the diagnosis of AASs; (2) they were cross-sectional diagnostic

studies; (3) prospective or retrospective enrollment was based on one or more AAS-compatible symptoms amongst chest pain, abdominal pain, back pain, syncope, perfusion deficit; (4) the ADD-RS was calculated; (5) the DD level was measured; (6) the diagnosis was confirmed or excluded with satisfactory criteria (advanced imaging with CTA, transesophageal echocardiography, magnetic resonance angiography, aortography, surgery or autopsy; in alternative, clinical case adjudication based on clinical data review and/or follow-up data); (7) absolute numbers of true positive, true negative, false positive, and false negative were reported or could be derived.

Studies were excluded if the design was case-control or case series due to high potential biases and the impossibility to calculate pre-test probability.[22] Conference abstracts were excluded because they are not peer-reviewed, the results may not be final and insufficient detail is provided for quality assessment.

Data extraction and analysis

The reporting of this systematic review follows the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA checklist, provided as *supplementary tables 2~~–3~~*).[21] Two reviewers extracted data independently from the selected articles. The extracted data included: first author, date of publication, study period, number of study sites, study setting, study design, inclusion/exclusion criteria, number of participants analyzed and excluded, D-dimer assay used, D-dimer reference range, reference standard used. They also extracted the study population characteristics (age, gender, time from symptom onset to evaluation), ADD-RS distribution, D-dimer level, AAS subtype, alternative final diagnoses made for patients without AASs, reporting the absolute number of true positive (TP), true negative (TN), false positive (FP), false negative (FN).

Two investigators independently assessed the quality of study design using the QUADAS-2 tool and the quality of reporting using the STARD tool.[23,24] QUADAS-2 assessment was done in compliance with the original background document.[23] For the domain "patient selection", we identified a high risk of bias if the sample was not consecutive, if the study wasn't done in the ED, if symptoms leading to patient inclusion did not include at least chest pain (representing the most common presenting symptom of AASs) and if patient enrollment was based on results of D-dimer or advanced imaging and not on clinical presentation. For the domain "index test", we identified a high risk of bias if the threshold of the index test wasn't prespecified or

if the result of the index test was interpreted after applying the reference standard. For the domain "reference standard", a high risk of bias was identified if patients were not subjected to advanced aortic imaging (CTA, transesophageal echocardiography, magnetic resonance angiography or aortography), surgery or autopsy. For patients not subjected to advanced imaging, surgery or autopsy, case adjudication based on independent clinical data review and/or follow-up data was considered satisfactory. For the domain "flow and timing", a high risk of bias was identified if studies included a significant (>5%) proportion of patients evaluated >14 days after symptom onset. Agreement between the reviewers was assessed with Cohen's k statistic. Types of diagnostic bias and anticipated skews in observed sensitivity/specificity were evaluated according to Kohn *et al*.[22]

Based on clinical reasoning and previous evidence, we planned to analyze the DD test results based on two different cutoffs: 500 ng/mL and an age-adjusted cutoff (DD_{adj}).[18] For the latter, the DD result was interpreted as follows: in patients younger than 50 years, an AAS was excluded in those with a DD value lower than 500 ng/mL. In patients aged 50 years or older, the DD test result was considered negative in those with a DD value lower than their age multiplied by 10. Briefly, DD_{adj} (ng/mL) was calculated as: age (years) × 10 ng/mL (with a minimum of 500 ng/mL).[15] To conduct an individual patient-level meta-analysis, the authors of all the selected studies were contacted to obtain missing data. For each study, a database was obtained reporting for each included patient, the age in years, the ADD-RS, the absolute DD level and the final diagnosis.

In the meta-analysis, we analyzed the performance of the following integrated strategies for diagnostic rule-out of AASs (*i.e.* if string satisfied, rule-out AASs): (1) ADD-RS≤1 and DD<500 ng/mL; (2) ADD-RS≤1 and $DD<DD_{age\text{-}adj}$; (3) ADD-RS=0 and DD<500 ng/mL; (4) ADD-RS=0 and $DD<DD_{age\text{-}adj}$. We built 2 x 2 contingency tables for each diagnostic strategy using the number of true positive (TP), false positive (FP), false negative (FN), true negative (TN). For negative likelihood ratio (LR) values of strategies with a sensitivity of 100%, contingency tables with zero value were handled by adding a 0.5 continuity correction and the 95% CI was estimated using a bootstrapping approach.[25] The failure rate was calculated as FN / (FN+TN), *i.e.* number of patients with AASs satisfying rule-out criteria divided by the total number of patients satisfying rule-out criteria.[26] The rule-out efficiency was calculated as (TN+FN) / (TP+FP+TN+FN), *i.e.* number of patients ruled-out by each integrated strategy divided by total number of patients tested. Heterogeneity was determined using

the Higgins' I^2. For variables showing non-significant heterogeneity, we calculated pooled values using fixed or random models as appropriate, based on inter- and intra-study variability.

The Pauker and Kassirer decision threshold model was applied to calculate two theoretical thresholds: a testing threshold (*i.e.* the probability of AAS at which there is no difference between performing the test and withholding the treatment) and a test-treatment threshold (*i.e.* the probability of AAS at which there is no difference between performing the test and administering the treatment).[27]

Statistical analysis was carried-out using Stata 13.1 (Stata Corp, College Station, Texas).

RESULTS

Literature search

Within 680 studies identified by the systematic database research, 12 studies were selected for full-text review (*figure 1*) and 4 studies met all the inclusion criteria.[16,28-30] Three studies were designed to investigate the diagnostic test characteristics of ADD-RS plus DD<500 ng/mL,[28-30] and one investigated the integration of ADD-RS with DD<$DD_{age-adj}$.[16] The study characteristics are summarized in *table 2* and the final diagnoses of the participant patients are detailed in *supplementary table ~~4~~3*. The case-mix of AASs was similar amongst studies, with acute aortic dissection representing the most frequent subtype and intramural aortic hematoma or penetrating aortic ulcer accounting for most of the other cases. Some specificities were found in the study by Kotani *et al*., which included a larger number of patients with complicated aneurysms (ruptured/with impending-rupture or infectious).[16] This study also reflects the higher prevalence of intramural hematomas in Japan and Asia. The pooled prevalence of AASs (mean ± SD) across the four studies was 18.0% ± 5.3%, which is substantially higher than reported in North American ED series, but also substantially lower than in most diagnostic biomarker studies, including the first key prospective multicenter study of D-dimer.[9,31-33] The pooled prevalence of "classic" acute aortic dissection was 67.3%, of intramural aortic hematoma was 18.7% and of penetrating aortic ulcer was 6.8%. A higher observed prevalence of AASs than in general ED practice could lead to spectrum bias (falsely raising sensitivity ~~and specificity~~).

The study by Nazerian *et al*. (2018) was the only prospective multicenter study. Its primary aim was to define the failure rate of a diagnostic rule-out strategy integrating ADD-RS (=0 or ≤1) with DD<500

ng/mL.[30] A secondary analysis applying $DD_{age-adj}$ has also been published.[17] In this trial, the gold standards for case-adjudication were conclusive aortic imaging (by CTA, transesophageal echocardiography or magnetic resonance angiography), surgery or autopsy. Enrollment preceded final decision on aortic imaging, and patients who were not subjected to any of these gold standards during the ED visit were subjected to 14-day follow-up. Patients or family members were interviewed by telephone with a structured questionnaire or underwent an outpatient visit after ED discharge, evaluating diagnosis of any aortic disease, subsequent ED visits, hospital admission and death. Reviewers defining case adjudication had access to hospital charts and ~~dismissal~~discharge documents. Gold standard imaging was obtained during the index visit in 45% of patients, and during follow-up in 1.6% of patients. Patients dismissed from the ED and with a negative follow-up were 34.3%, potentially leading to differential verification bias, with a decrease in observed sensitivity and specificity. During follow-up, 2 patients were lost and 3 cases of AASs were diagnosed. Three patients died without advanced imaging or surgery. They all had a positive DD test result and therefore could not be regarded as potential FN cases.

The studies by Nazerian *et al*. (2014), Gorla *et al*. and Kotani *et al*. were retrospective.[16,28,29] In these studies, an exact time definition of symptoms triggering ~~enrolment~~enrollment was not reported by the authors. This raises concern about the potential inclusion of patients with non-acute symptoms, in whom the chance of FN cases is higher. In the study by Nazerian *et al*. (2014), data were obtained from a registry of ED patients undergoing advanced aortic imaging for clinically suspected AAS.[28] The study cohort largely overlapped with another study from the same groups which focused on validation of the ADD-RS *per se*.[34] For 29% of enrolled participants, a DD test result was not available, leading to patient exclusion. This could introduce partial verification bias (which could raise sensitivity), but the characteristics of the patients in the included and excluded groups were similar. A *post*-hoc analysis showed that only 17 patients (1.6%) presented with history of pain >14 days; 4 of them had an AAS. One of these patients (symptoms for 15 days), had a normal D-dimer. However, exact time data was missing for 39.2% of the enrolled patients.

In Gorla *et al*.,[29] patients were enrolled if they were admitted to the ED for chest pain and if they were subjected to a DD assay. These criteria could bias against atypical presentations not involving chest pain. The clinical judgment of the physician ordering DD was not recorded. Hence, PE and not AAS could have represented the chief differential diagnosis in some patients. Indeed, the rate of PE in this cohort was 14.5%,

significantly higher than in the other studies. The authors declare that in study patients, CTA was used *per* guidelines, based on clinical judgment and on DD test result. All diagnoses of AASs were confirmed by advanced aortic imaging. Since the actual number of patients subjected to advanced imaging is unknown and clinical case adjudication was not based on a pre-specified follow-up, observed sensitivity could be raised due to differential verification bias.

In the study by Kotani *et al*.,[16] patients were also enrolled if they presented with acute chest pain and if they received a DD assay. The exact time interval from symptom onset to sampling was not presented, and the DD assay was used *per* a pre-specified hospital protocol not detailed in the manuscript. The analysis was conducted only on patients admitted to hospital after the ED visit, while patients dismissed from the ED were excluded. This could lead to spectrum bias, raising sensitivity in the enrolled sample. Restriction to admitted patients potentially biases towards a more clinically severe population, while rule-out strategies ideally apply to patients in whom early ED ~~dismissal~~discharge represents a meaningful option. However, the final prevalence of AASs was 13.9%, indicating adequate representation of low-probability patients. Additional exclusion criteria were ST elevation on ECG and hemodynamic instability. Both criteria are in line with ESC recommendations, as patients with these clinical characteristics are not amenable to rule-out criteria.[11] The DD assay was interpreted using the $DD_{age\text{-}adj}$ cutoff. As in Gorla *et al*., ~~enrolment~~enrollment criteria focused on chest pain, excluding alternative clinical presentations and likely included patients with a clinical suspicion of PE and not only of AASs. However, the prevalence of PE was generally low (3.8%), while the prevalence of acute coronary syndromes was the highest, indicating potential bias towards coronary artery disease.

Quality assessment

The quality assessment (QA) conducted using the QUADAS-2 is shown in *table 3* and in *supplementary figure 1*. For only one study, the judgment was "low" in all 7 domains, indicating an overall low risk of bias and concern regarding applicability.[30] In one study, the judgment was "low" in 3 of 7 domains.[29]

The quality of reporting of the included studies, analyzed according to the STARD 2015 statement, is detailed in *supplementary table ~~5~~4*. Most studies showed suboptimal quality regarding type of sample enrollment, how missing data on the index test and reference standard were handled, sample size calculation,

whether any clinical intervention was done between the index test and the reference standard, study registration and accessibility of the full protocol. The agreement between the reviewers for components of the study quality assessment tools was good (κ=0.67, 95% CI 0.54 – 0.80).[35]

Meta-analysis

A total of 3804 patients were included in the meta-analysis, including 675 (17.7%) with AASs. To evaluate strategies integrating either the 500 ng/mL or the $DD_{age\text{-}adj}$ cutoff, individual patient-level data were used. Contingency tables and coupled forest plots were obtained (*figure 2*). For all strategies, statistical heterogeneity was negligible for sensitivity (I^2=0%) and significant for specificity values. Subanalyses excluding patients with ADD-RS=0, shown in *supplementary tables 5-6-~~7~~*, indicated that results were not substantially affected by inclusion of patients at lowest pre-test probability of AASs. Negative and positive likelihood ratio (LR) values of the diagnostic strategies are shown in *figure 3*. Heterogeneity was negligible for the negative LR (I^2=0%) and significant for the positive LR values of the diagnostic strategies.

Forest plots of failure rate and efficiency values are shown in *figures 4-5*. Failure rate values had low to moderate heterogeneity for ADD-RS=0 and DD<500 ng/mL (I^2=38.1%), ADD-RS=0 and DD<$DD_{age\text{-}adj}$ (I^2=28%), ADD-RS≤1 and DD<500 ng/mL (I^2=39%); heterogeneity was significant for ADD-RS≤1 plus DD<$DD_{age\text{-}adj}$ (I^2=84.4%). Efficiency values had significant heterogeneity for all diagnostic strategies.

Pooled estimates of diagnostic variables underlying diagnostic rule-out (sensitivity, negative LR and failure rate) and showing non-significant heterogeneity across studies are summarized in *table 4*. Diagnostic variables showing high heterogeneity were not to reported, as limited inference on pooled values can be done. For ADD-RS=0 and DD<500 ng/mL, pooled sensitivity was 99.9% (99.3–100%), negative LR 0.032 (0–0.086) and failure rate 0.1% (0–0.3%). For ADD=0 and DD<$DD_{age\text{-}adj}$ sensitivity was 99.9% (99.3–100%), negative LR 0.027 (0–0.081) and failure rate 0.1% (0–0.2%). For ADD-RS≤1 and DD<500 ng/mL, sensitivity was 98.9% (97.9–99.9%), negative LR 0.025 (0.001–0.049) and failure rate 0.6% (0.2–0.9%). For ADD-RS≤1 and DD<$DD_{age\text{-}adj}$, sensitivity was 97.6% (96.3–98.9%) and negative LR 0.048 (0.022–0.074). For this strategy, pooled failure was not computed due to significant heterogeneity.

Test-treatment threshold

Test-treatment thresholds were calculated for diagnostic strategies including DD<500 ng/ml (*supplementary figure 2*). According to this model, the ADD-RS=0 and DD<500 ng/mL strategy should be performed if the clinical probability of AASs is between 1.7% and 23.2%, while the ADD-RS≤1 and DD<500 ng/mL strategy should be performed when the pretest probability is between 1.1% and 44.8%.

DISCUSSION

We provide a systematic review and summary of studies assessing integration of ADD-RS with DD for diagnosis of AASs. Only four papers satisfied the pre-defined inclusion criteria, underlying the relative paucity of data. However, the total number of included patients was substantial (n=3804). All studies post-dated the latest guidelines of the American Heart Association and the European Society of Cardiology, and only one (Nazerian *et al*. 2014) was cited in the latest clinical policy of the American College of Emergency Physicians.[11,12,36] One was a prospective multicenter trial, while the other 3 were retrospective studies. All were performed in the ED and mostly involved patients with chest pain, but inclusion criteria partly differed. This key limit reflects the absence of a standard definition of patients suspected of having AASs and amenable to rule-out strategies. Therefore, methodological and clinical heterogeneity between available studies mandate caution in efforts to pool and summarize data.

Significant statistical heterogeneity was found for specificity, positive LR and efficiency. This likely reflects differences in the clinical case-mix of study cohorts. For these variables, data pooling could be misleading and was therefore omitted. AASs cases were instead homogenous across studies, thus leading to negligible statistical heterogeneity for sensitivity and negative LR values and allowing meaningful data pooling for these variables. Also in a previous meta-analysis of high-quality studies (which also included Nazerian *et al*. 2014), the heterogeneity was low for sensitivity and negative LR, and substantial for specificity and positive LR.[14]

Acquisition of primary data allowed us to evaluate diagnostic strategies incorporating also $DD_{age\text{-}adj}$, already in use for PE rule-out. When using ADD-RS=0, $DD_{age\text{-}adj}$ provided pooled sensitivity and negative LR values similar to those of the "classical" 500 ng/mL cutoff. Instead, when using ADD-RS≤1, DD<500 ng/mL outperformed DD<$DD_{age\text{-}adj}$ in terms of pooled sensitivity and negative LR. This data suggests that DD<$DD_{age\text{-}}$

$_{adj}$ could be evaluated in further studies only if the pre-test probability is presumed to be very low. $DD_{age\text{-}adj}$ might provide increased specificity over 500 ng/mL, but the statistical heterogeneity found across studies does not allow any conclusion.

Consensus is lacking on what should reproducibly define a clinical suspicion of AASs. Hence, differences between physicians and centers can be profound. In North American retrospective series of patients undergoing CTA for suspected AAS, the prevalence of AASs was ~3%.[8,37] In a vast out-of-hospital study evaluating the ADD-RS in non-traumatic emergencies, the prevalence of AASs was 0.9%.[38] In the studies reviewed herein, the prevalence of AASs was 13 to 23%. Application of rule-out strategies to patient populations at lower pre-test probability of AASs is expected to result in lower failure rates, with a trade-off in efficiency.

Caution is needed when considering application of ADD-RS and D-dimer based strategies in clinical practice. First, ADD-RS, a decision rule derived from a retrospective register of AASs, has low specificity.[37] In addition, ADD-RS derivation methods have not been published, and it is currently unknown whether use of the ADD-RS provides any advantage in terms of diagnostic accuracy and of CTA ordering, as compared to clinical gestalt.[39,40] In the future, focused ED-centered studies may provide alternative and more specific probability assessment tools. Second, ~~also~~ D-dimer also lacks specificity. Therefore, indiscriminate application of ADD-RS and D-dimer to unselected ED patients with AAS-compatible symptom(s) would paradoxically increase the number of CTA ordered. Such slippery slope must be avoided.[41]

Based on previous data, in terms of specificity, we speculate that the ADD-RS/D-dimer rule-out pathway could best apply to stable patients with ADD-RS=1 owing to clinical manifestations providing *per se* higher specificity (*i.e.* pulse deficit, neurological deficit, aortic valve insufficiency).[42] Caution is needed in patients with hypotension, which also potentially defines clinical instability and might prompt towards a fast-track for advanced imaging irrespective of D-dimer test results. However, in clinical practice, most cases with ADD-RS=1 will be driven by pain features (severe, sudden, ripping pain), providing higher sensitivity but lower specificity. To maximize benefits, a pragmatic approach could be to request D-dimer only after three-dimensional evaluation of clinical history, physical examination, first-line imaging and blood test results, in patients still lacking a clear alternative diagnosis, or in whom rule-out of AASs is considered imperative for decision on hospital admission versus discharge or administration of anticoagulant/antiplatelet therapies,

which could be harmful in presence of an AAS.

LIMITATIONS

Only one study (49% of patients) was judged to provide a low risk of bias/applicability concerns.[30] Two studies (42% of patients)[16,28] had issues in one of the QUADAS domains, and one study (10% of patients) had a generally lower quality profile.[29] In one study, the case-mix of AASs slightly differed, with fewer cases of acute aortic dissections and higher prevalence of the other forms.[16] Overall, the potential bias types most frequently encountered were: (1) partial verification bias, due to patients excluded because discharged from the ED or due to unavailable DD test result (leading to potential upward skew in sensitivity and downward skew in specificity), and (2) differential verification bias, due to inclusion of patients subjected to clinical follow-up without advanced aortic imaging (leading to potential downward skew in sensitivity and specificity). The accuracy of DD for diagnosis of AASs may also slightly differ amongst subtypes, with higher risk of false negative cases in patients with intramural hematomas and focal dissections.[43,44] Therefore, methodological and clinical heterogeneity between available studies mandate caution in data pooling and summarization.

A key issue affecting two studies (24% of patients) is that the authors selected patients with chest pain and a DD test result, potentially also introducing ~~also~~ individuals with suspected PE.[16,29] In clinical terms, this aspect may be secondary, because both PE and AASs are typically considered in differential diagnosis, share DD as the key biomarker and require CTA for conclusive diagnosis. A suspicion of AAS by the attending physicians was clearly defined in two studies (76% of patients) led by the same primary investigators.[28,30] This might limit external validity.

With the exception of the ADvISED trial, there was general uncertainty about the timing of the index test. Hence, a minority of patients with symptoms dating >14 days were possibly enrolled, including few cases of AASs in their subacute or chronic phase. Since D-dimer levels tend to ~~reduce~~decrease over time after development of AASs, this is expected to increase the number of patients with AASs presenting as FN (differential verification bias, with potential downward skew in estimates of sensitivity and failure rate).[45]

CONCLUSIONS

Only 4 studies have evaluated integration of ADD-RS with DD for diagnosis of AASs, with methodological differences that must be carefully considered. However, the total number of included patients is reasonably large (n=3804), and negligible heterogeneity was found for sensitivity and negative LR values. Available studies consistently show that ADD-RS=0 or ≤1 plus DD<500 ng/mL are highly sensitive diagnostic strategies and support their reliability for rule-out of AASs. For $DD_{age\text{-}adj}$, available data appear largely preliminary and further studies are required. Nonetheless, further prospective trials, especially in low prevalence populations, are needed to confirm the results of this meta-analysis.

REFERENCES

1. Bossone E, LaBounty TM, Eagle KA. Acute aortic syndromes: diagnosis and management, an update. Eur Heart J 2018;39:739-49d.

2. Goodacre S, Cross E, Arnold J, Angelini K, Capewell S, Nicholl J. The health care burden of acute chest pain. Heart 2005;91:229-30.

3. Alshekhlee A, Shen WK, Mackall J, Chelimsky TC. Incidence and mortality rates of syncope in the United States. Am J Med 2009;122:181-8.

4. Amsterdam EA, Kirk JD, Bluemke DA, et al. Testing of low-risk patients presenting to the emergency department with chest pain: a scientific statement from the American Heart Association. Circulation 2010;122:1756-76.

5. Owens PL, Barrett ML, Gibson TB, Andrews RM, Weinick RM, Mutter RL. Emergency department care in the United States: a profile of national data sources. Ann Emerg Med 2010;56:150-65.

6. Klompas M. Does this patient have an acute thoracic aortic dissection? JAMA 2002;287:2262-72.

7. Hansen MS, Nogareda GJ, Hutchison SJ. Frequency of and inappropriate treatment of misdiagnosis of acute aortic dissection. Am J Cardiol 2007;99:852-6.

8. Lovy AJ, Bellin E, Levsky JM, Esses D, Haramati LB. Preliminary development of a clinical decision rule for acute aortic syndromes. Am J Emerg Med 2013;31:1546-50.

9. Ohle R, Anjum O, Bleeker H, Wells G, Perry JJ. Variation in emergency department use of computed tomography for investigation of acute aortic dissection. Emerg Radiol 2018;25:293-8.

10. Rogers AM, Hermann LK, Booher AM, et al. Sensitivity of the aortic dissection detection risk score, a novel guideline-based tool for identification of acute aortic dissection at initial presentation: results from the international registry of acute aortic dissection. Circulation 2011;123:2213-8.

11. Erbel R, Aboyans V, Boileau C, et al. 2014 ESC Guidelines on the diagnosis and treatment of aortic diseases: Document covering acute and chronic aortic diseases of the thoracic and abdominal aorta of the adult. The Task Force for the Diagnosis and Treatment of Aortic Diseases of the European Society of Cardiology (ESC). Eur Heart J 2014;35:2873-926.

12. Diercks DB, Promes SB, Schuur JD, Shah K, Valente JH, Cantrill SV. Clinical policy: critical issues in the evaluation and management of adult patients with suspected acute nontraumatic thoracic aortic dissection. Ann Emerg Med 2015;65:32-42 e12.

13. Konstantinides SV, Torbicki A, Agnelli G, et al. 2014 ESC guidelines on the diagnosis and management of acute pulmonary embolism. Eur Heart J 2014;35:3033-69, 69a-69k.

14. Asha SE, Miers JW. A Systematic Review and Meta-analysis of D-dimer as a Rule-out Test for Suspected Acute Aortic Dissection. Ann Emerg Med 2015;66:368-78.

15. Righini M, Van Es J, Den Exter PL, et al. Age-adjusted D-dimer cutoff levels to rule out pulmonary embolism: the ADJUST-PE study. JAMA 2014;311:1117-24.

16. Kotani Y, Toyofuku M, Tamura T, et al. Validation of the diagnostic utility of D-dimer measurement in patients with acute aortic syndrome. Eur Heart J Acute Cardiovasc Care 2017;6:223-31.

17. Morello F, Mueller C, Soeiro AM, et al. Response by Morello et al to Letters Regarding Article, "Diagnostic Accuracy of the Aortic Dissection Detection Risk Score Plus D-Dimer for Acute Aortic Syndromes: The ADvISED Prospective Multicenter Study". Circulation 2018;138:448-9.

18. Suzuki T, Eagle KA. Biomarker-Assisted Diagnosis of Acute Aortic Dissection. Circulation 2018;137:270-2.

19. Baez AA, Cochon L. Improved rule-out diagnostic gain with a combined aortic dissection detection risk score and D-dimer Bayesian decision support scheme. J Crit Care 2017;37:56-9.

20. Deeks JJ, Bossuyt PM, Gatsonis C. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0. 2013.

21. McInnes MDF, Moher D, Thombs BD, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA 2018;319:388-96.

22. Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. Acad Emerg Med 2013;20:1194-206.

23. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529-36.

24. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.

25. Marill KA, Chang Y, Wong KF, Friedman AB. Estimating negative likelihood ratio confidence when test sensitivity is 100%: A bootstrapping approach. Stat Methods Med Res 2017;26:1936-48.

26. van Es N, van der Hulle T, van Es J, et al. Wells Rule and d-Dimer Testing to Rule Out Pulmonary Embolism: A Systematic Review and Individual-Patient Data Meta-analysis. Ann Intern Med 2016;165:253-61.

27. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. N Engl J Med 1980;302:1109-17.

28. Nazerian P, Morello F, Vanni S, et al. Combined use of aortic dissection detection risk score and D-dimer in the diagnostic workup of suspected acute aortic dissection. Int J Cardiol 2014;175:78-82.

29. Gorla R, Erbel R, Kahlert P, et al. Accuracy of a diagnostic strategy combining aortic dissection detection risk score and D-dimer levels in patients with suspected acute aortic syndrome. Eur Heart J Acute Cardiovasc Care 2017;6:371-8.

30. Nazerian P, Mueller C, Soeiro AM, et al. Diagnostic Accuracy of the Aortic Dissection Detection Risk Score Plus D-Dimer for Acute Aortic Syndromes: The ADvISED Prospective Multicenter Study. Circulation 2018;137:250-8.

31. Suzuki T, Distante A, Zizza A, et al. Diagnosis of acute aortic dissection by D-dimer: the International Registry of Acute Aortic Dissection Substudy on Biomarkers (IRAD-Bio) experience. Circulation 2009;119:2702-7.

32. Taylor RA, Iyer NS. A decision analysis to determine a testing threshold for computed tomographic angiography and d-dimer in the evaluation of aortic dissection. Am J Emerg Med 2013;31:1047-55.

33. Wang Y, Tan X, Gao H, et al. Magnitude of Soluble ST2 as a Novel Biomarker for Acute Aortic Dissection. Circulation 2018;137:259-69.

34. Nazerian P, Giachino F, Vanni S, et al. Diagnostic performance of the aortic dissection detection risk score in patients with suspected acute aortic dissection. Eur Heart J Acute Cardiovasc Care 2014;3:373-81.

35. Byrt T. How good is that agreement? Epidemiology 1996;7:561.

36. Hiratzka LF, Bakris GL, Beckman JA, et al. 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM guidelines for the diagnosis and management of patients with Thoracic Aortic Disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, American College of Radiology, American Stroke Association, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society of Interventional Radiology, Society of Thoracic Surgeons, and Society for Vascular Medicine. Circulation 2010;121:e266-369.

37. Ohle R, Anjum O, Bleeker H, McIsaac S. What Is the Specificity of the Aortic Dissection Detection Risk Score in a Low-prevalence Population? Acad Emerg Med 2019;26:632-8.

38. Yamashita A, Maeda T, Kita Y, et al. The impact of prehospital assessment and EMS transport of acute aortic syndrome patients. Am J Emerg Med 2018;36:1188-94.

39. Hill JM, Murphy TG, Fermann GJ. Aortic Dissection Detection Risk Score: A Clinical Decision Rule That Needs Some Parenting. Acad Emerg Med 2019;26:695-7.

40. Smith LM, Miller CD. Acute Aortic Dissection: Is There Something Better than Physician Gestalt? Acad Emerg Med 2018;25:464-6.

41. Fatovich DM. The inverted U curve and emergency medicine: Overdiagnosis and the law of unintended consequences. Emerg Med Australas 2016;28:480-2.

42. Ohle R, Kareemi HK, Wells G, Perry JJ. Clinical Examination for Acute Aortic Dissection: A Systematic Review and Meta-analysis. Acad Emerg Med 2018;25:397-412.

43. Ohlmann P, Faure A, Morel O, et al. Lower circulating Sta-Liatest D-Di levels in patients with aortic intramural hematoma compared with classical aortic dissection. Crit Care Med 2009;37:899-901.

44. Gorla R, Erbel R, Kahlert P, et al. Diagnostic role and prognostic implications of D-dimer in different classes of acute aortic syndromes. Eur Heart J Acute Cardiovasc Care 2017;6:379-88.

45. Eggebrecht H, Naber CK, Bruch C, et al. Value of plasma fibrin D-dimers for detection of acute aortic dissection. J Am Coll Cardiol 2004;44:804-9.

FIGURE LEGENDS

Figure 1. PRISMA flow diagram of study search and selection.

Figure 2. Contingency tables and coupled forest plots of sensitivity and specificity values. Heterogeneity was determined using the Higgins' I^2.

Figure 3. Forest plots of the negative and positive likelihood ratio values. Heterogeneity was determined using the Higgins' I^2.

Figure 4. Forest plots of the failure rate values. Heterogeneity was determined using the Higgins' I^2.

Figure 5. Forest plots of the efficiency values. Heterogeneity was determined using the Higgins' I^2.

TABLES

Table 1. Detailed database search strategies.

Literature database	Search query
MEDLINE	((((("Aneurysm, Dissecting"[Mesh]) AND "Fibrin Fibrinogen Degradation Products"[Mesh]) OR (acute aortic syndrome AND D-dimer))) OR "Aortic Dissection Detection Risk Score")
EMBASE	(('acute aortic syndrome'/exp OR 'acute aortic syndrome' OR 'aortic dissection'/exp OR 'aortic dissection') AND ('d dimer'/exp OR 'd dimer') OR 'aortic dissection detection risk score') NOT 'conference abstract':it NOT review:it NOT letter:it
Web Of Science	TOPIC: (("acute aortic syndrome" OR "aortic dissection" OR "dissecting aneurysm" OR "Aortic Dissection Detection Risk Score") AND ("D-dimer" OR "Fibrin Degradation Product"))

Table 2. Characteristics of included studies.

	Nazerian, 2014	**Gorla, 2017**	**Kotani, 2017**	**Nazerian, 2018**
Study period	01/2008 to 03/2013	01/2001 to 05/2014	01/2011 to 04/2014	01/2014 to 12/2016
N. of study sites	2	1	1	6
Setting country hospital department	 IT Large referral ED	 GE NR ED	 JA Large referral ED	 BR, SW, GE, IT Large referral ED
Participants, N (% of enrolled)	1035 (71%)	376 (100%)	545 (61.4%)	1848 (99.9%)
Participants excluded for unavailable index test, N (%)	420 (29%)	0 (0%)	66 (6.9%)	48 (2.5%)
AASs, N (% enroll.) AD, N (% AAS) IMH, N (% AAS) PAU, N (% AAS) other, N (% AAS)	233 (22.5%) 199 (85.4%) 31 (13.3%) 3 (1.3%) 0	85 (22.6%) 61 (71.8%) 11 (12.9%) 13 (15.3%) 0	123 (13.9%) 47 (38.2%) 42 (34.1%) 8 (6.5%) 26 (21.1%)[#]	241 (13%) 178 (73.9%) 35 (14.5%) 10 (4.1%) 18 (7.5%)[¶]
ADD-RS, N (N, % with AAS) 0 1 2-3	 322 (19, 5.9%) 508 (133, 26.2%) 205 (81, 39.5%)	 189 (1, 0.5%) 130 (30, 23.1%) 57 (54, 94.7%)	 75 (4, 5.3%)[§] 399 (88, 22.1%)[§] 71 (24, 33.8%)[§]	 437 (12, 2.7%) 1070 (96, 9.0%) 341 (133, 39.0%)
Study design	Prospective enrollment, retrospective analysis	Retrospective	Retrospective	Prospective
Inclusion criteria	chest/back/ abdominal pain, syncope or perfusion deficit + alt-D not established + clinical suspicion leading to CTA	chest pain + D-dimer available at presentation	acute chest pain + admission to hospital + D-dimer available	chest/back/abdominal pain, syncope or perfusion deficit + clinical suspicion
Exclusion criteria	NR	NR	hemodynamic instability, STEMI, ED discharge, death	primary trauma, unwillingness or inadequacy to participate

			in ED, referral to other hospital	
Patient sampling	NR	NR	NR	Consecutive
Reference standard	CTA	unspecified advanced imaging study	CTA	CTA, TEE, MRA, surgery or autopsy; if unavailable, 14-day clinical follow-up
Age, y mean (SD)	67 (14%)	63 (12%)	70 (14%)	62 (12%)
Male, %	66%	61%	63.4%	62.3%
Duration of symptoms, hours	48 (7-96)*§	NR	82% <24 h	7.5 (2-30)*
D-dimer assay	HemosIL D-Dimer HS, STA®-Liatest® D-Di	Innovance® D-Dimer	Liatest D-dimer, Hexamate D-dimer	HemosIL D-Dimer HS, STA®-Liatest® D-Di, TriniLIA D-Dimer, Innovance® D-Dimer
D-dimer cutoff (ng/ml)	<500 ng/ml	≤500 ng/ml	if age ≤50y: < 500 ng/ml if age >50y: < (age x 10) ng/ml	<500 ng/ml
D-dimer, test char. sensitivity specificity	 98.3% 35.9%	 97.6% 63.2%	 96.0% 58.0%	 96.7% 64.0%

AAS: acute aortic syndrome; AD: aortic dissection; alt-D: alternative diagnosis; BR: Brazil; CTA: computed tomography angiography; GE: Germany; IMH: intramural aortic hematoma; IQR: interquartile range; IT: Italy; JA: Japan; MRA: magnetic resonance angiography; NR: not reported; PAU: penetrating aortic ulcer; SD: standard deviation; SW: Switzerland; TEE: transesophageal echocardiography.

*Values reported as median (IQR). §Original data provided by the authors for the present analysis and not included in the original manuscript. #Includes: ruptured aortic aneurysm (7.3%), impending rupture of aortic aneurysm (10.6%), infectious aortic aneurysm (3.2%). ¶Includes only spontaneous (non-traumatic) rupture of thoracic aorta.

Table 3. Assessment of study quality according to QUADAS-2.[23]

Study	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Nazerian, 2014	L	L	L	U	L	L	L
Gorla, 2017	H	L	U	U	L	L	U
Kotani, 2017	H	L	L	L	L	L	L
Nazerian, 2018	L	L	L	L	L	L	L

L =Low-risk, H =High-risk, U =Unclear

Table 4. Pooled estimates of diagnostic variables underlying diagnostic rule-out.

	Sensitivity (%, 95% CI)	**Negative LR (95% CI)**	**Failure rate (%, 95% CI)**
ADD-RS=0 and DD<500 ng/mL	99.9% (99.3–100%)	0.032 (0-0.086)	0.1% (0–0.3%) 1 in 1000 (333–∞)
I-squared, *p*	0%, 0.95	0%, 0.64	38.1%, 0.18
ADD-RS=0 and DD<$DD_{age\text{-}adj}$	99.9% (99.3–100%)	0.027 (0-0.081)	0.1% (0–0.2%) 1 in 1000 (500–∞)
I-squared, *p*	0%, 0.95	0%, 0.77	28%, 0.24
ADD-RS≤1 and DD<500 ng/mL	98.9% (97.9–99.9%)	0.025 (0.001–0.049)	0.6% (0.2–0.9%) 1 in 167 (111–500)
I-squared, *p*	0%, 0.91	0%, 0.98	39%, 0.19
ADD-RS≤1 and DD<$DD_{age\text{-}adj}$	97.6% (96.3–98.9%)	0.048 (0.022–0.074)	n.a.
I-squared, *p*	0%, 0.86	0%, 0.56	84.4%, <0.001

LR: likelihood ratio; n.a.=not applicable due to significant heterogeneity.

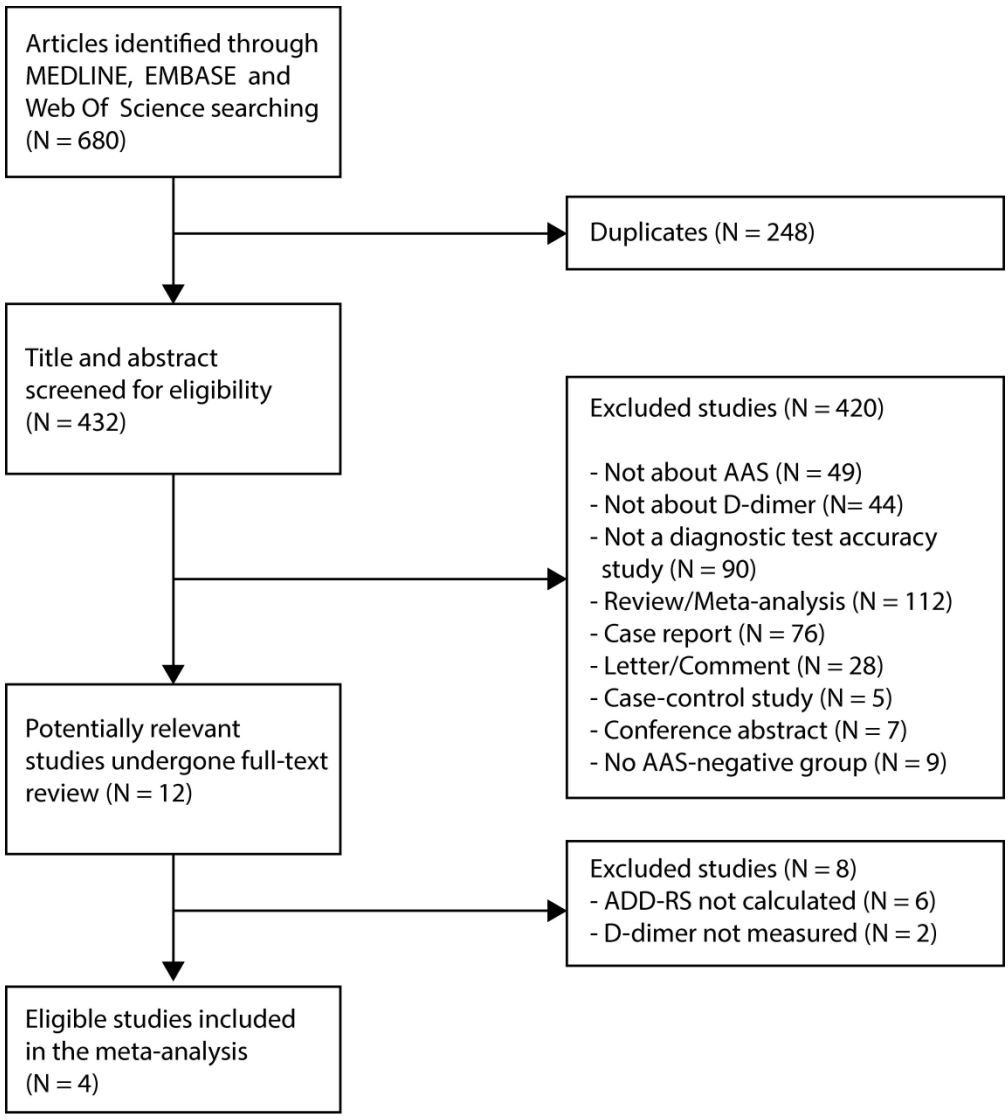


Figure 1. PRISMA flow diagram of study search and selection.

157x177mm (600 x 600 DPI)

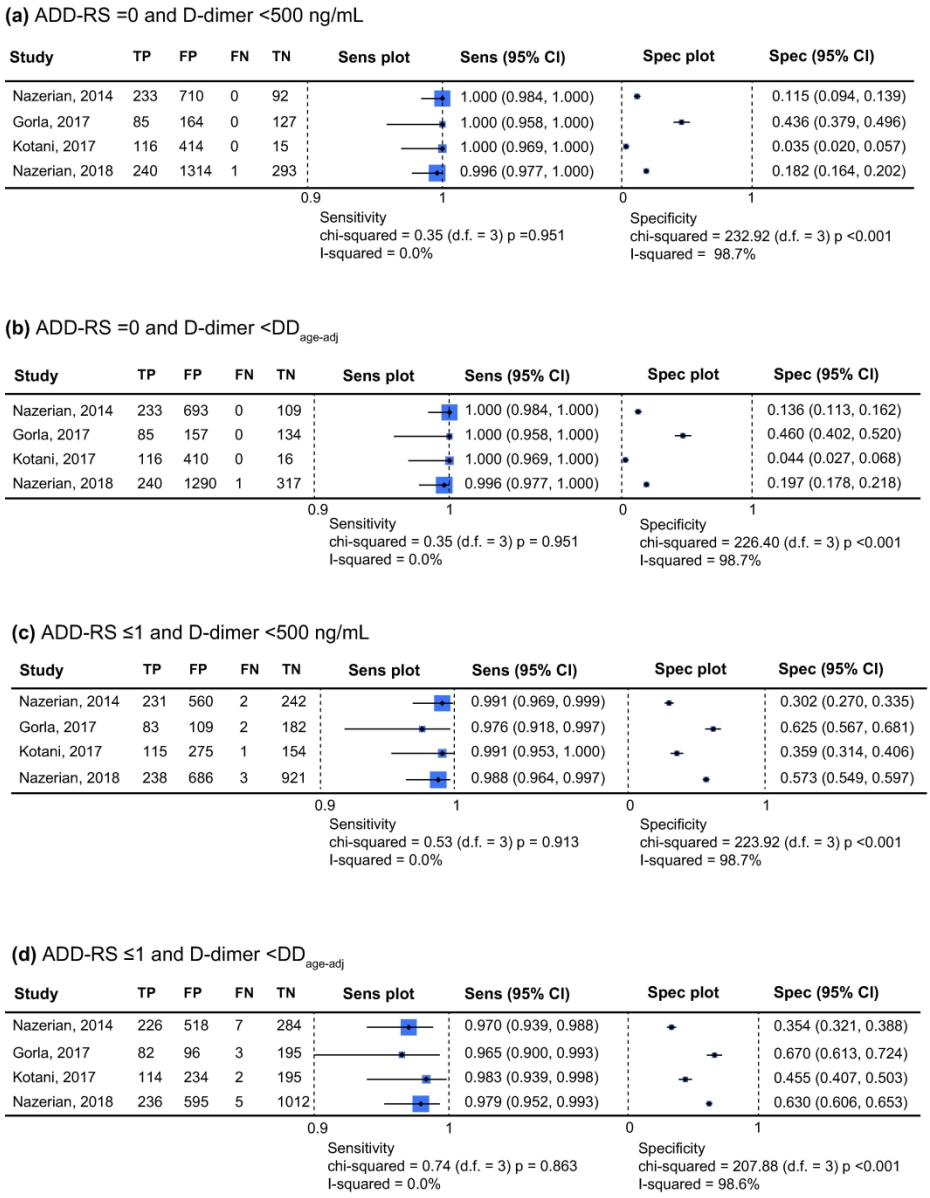


Figure 2. Contingency tables and coupled forest plots of sensitivity and specificity values. Heterogeneity was determined using the Higgins' I2.

176x227mm (600 x 600 DPI)

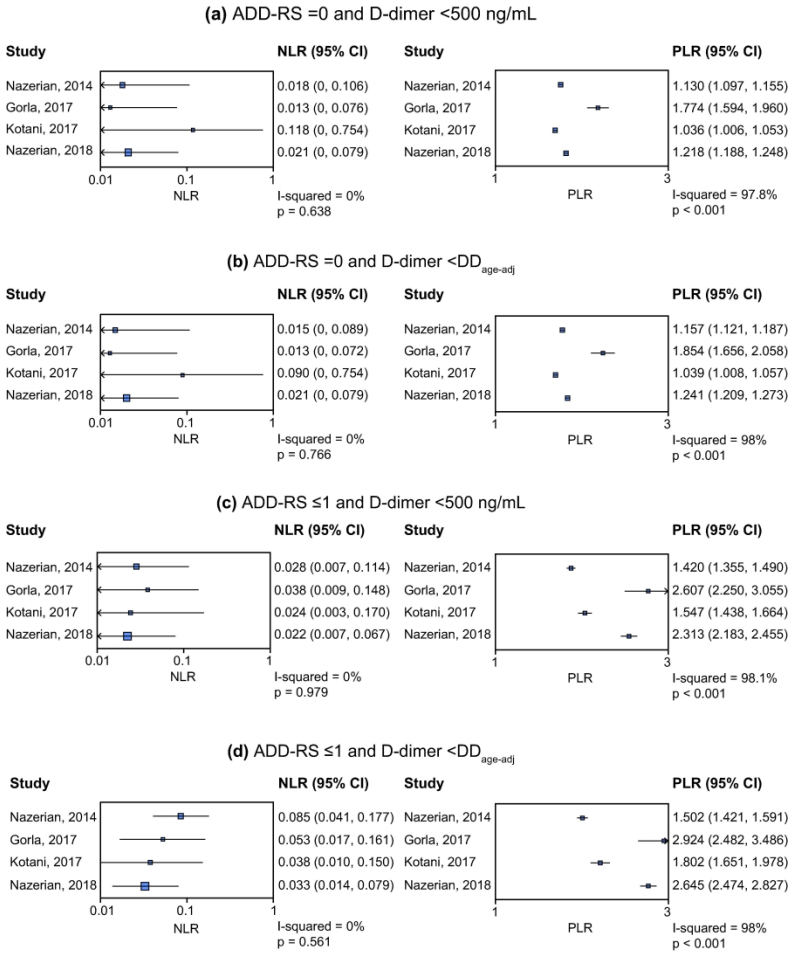


Figure 3. Forest plots of the negative and positive likelihood ratio values. Heterogeneity was determined using the Higgins' I2.

182x263mm (600 x 600 DPI)

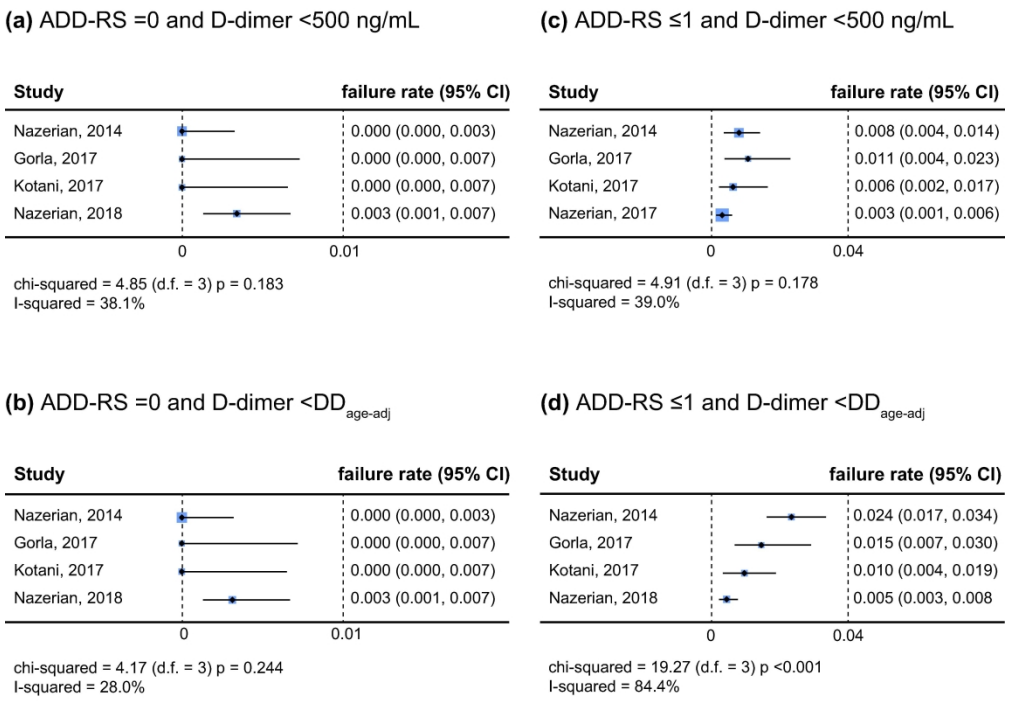


Figure 4. Forest plots of the failure rate values. Heterogeneity was determined using the Higgins' I2.

175x124mm (600 x 600 DPI)

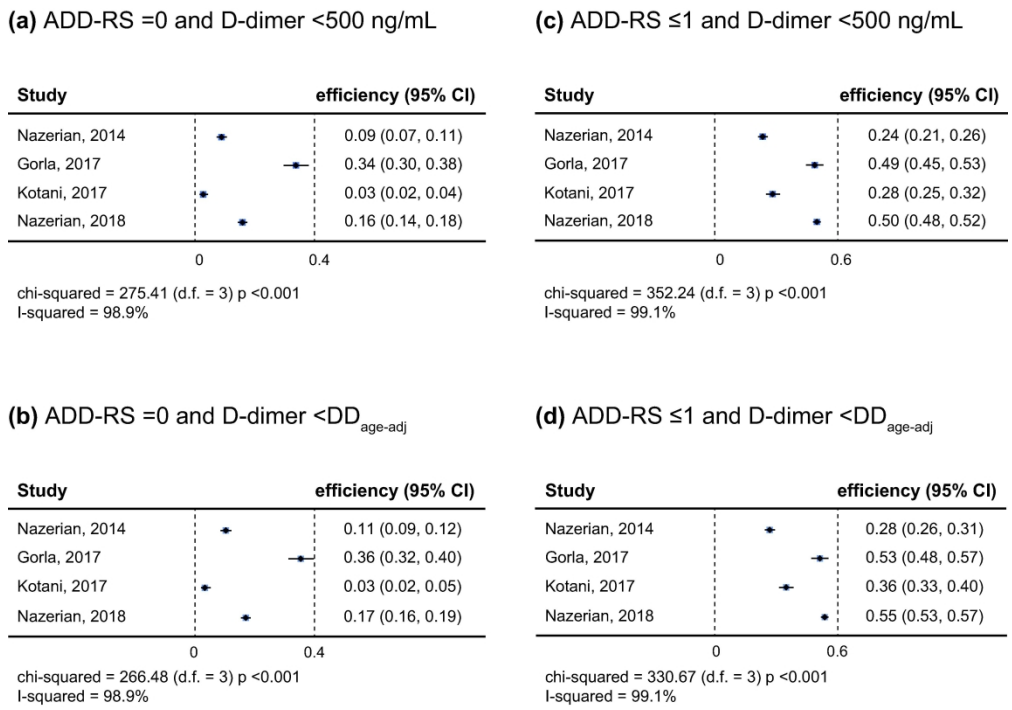


Figure 5. Forest plots of the efficiency values. Heterogeneity was determined using the Higgins' I2.

175x127mm (600 x 600 DPI)

Supplementary Data File

Systematic review of aortic dissection detection risk score plus D-dimer for diagnostic rule-out of suspected acute aortic syndromes

- Supplementary Data -

Summary

Supplementary Data File

Supplementary table 1. Aortic Dissection Detection Risk Score (ADD-RS).

High-risk conditions
• Marfan syndrome • Family history of aortic disease • Known aortic valve disease • Recent aortic manipulation • Known thoracic aortic aneurysm
High-risk pain features
Chest, back, or abdominal pain described as: • Abrupt in onset • Severe in intensity • Ripping or tearing in quality
High-risk exam features
• Pulse deficit or systolic BP differential • Focal neurologic deficit (with pain) • Murmur of aortic insufficiency (new, with pain) • Hypotension or shock state

Supplementary Data File

Supplementary table 2. Prisma-DTA checklist for abstracts, from McInnes MDF *et al*.

Section/topic	#	PRISMA-DTA for Abstracts Checklist item	Reported on page # of the manuscript
TITLE and PURPOSE			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Objectives	2	Indicate the research question, including components such as participants, index test, and target conditions.	1
METHODS			
Eligibility criteria	3	Include study characteristics used as criteria for eligibility.	-
Information sources	4	List the key databases searched and the search dates.	1 (no dates)
Risk of bias & applicability	5	Indicate the methods of assessing risk of bias and applicability.	1
Synthesis of results	A1	Indicate the methods for the data synthesis.	
RESULTS			
Included studies	6	Indicate the number and type of included studies and the participants and relevant characteristics of the studies (including the reference standard).	1
Synthesis of results	7	Include the results for the analysis of diagnostic accuracy, preferably indicating the number of studies and participants. Describe test accuracy including variability; if meta-analysis was done, include summary results and confidence intervals.	1
DISCUSSION			
Strengths and limitations	9	Provide a brief summary of the strengths and limitations of the evidence	1
Interpretation	10	Provide a general interpretation of the results and the important implications.	1
OTHER			
Funding	11	Indicate the primary source of funding for the review.	-
Registration	12	Provide the registration number and the registry name	-

Supplementary Data File

Supplementary table 3. Final diagnoses of patients in the included studies.

	Nazerian, 2014	Gorla, 2017	Kotani, 2017	Nazerian, 2018
AASs, n (%tot)	233 (22.5)	85 (22.6)	123 (13.9)	241 (13.0)
Stanford Type-A AD, N (%AAS)	148 (63.5)	33 (38.8)	27 (22.0)§	125 (51.9)
Stanford Type-B AD, N (%AAS)	51 (21.9)	28 (32.9)	20 (16.3)§	53 (22.0)
IMH, N (%AAS)	31 (13.33)	11 (12.9)	42 (34.1)	35 (14.5)
PAU, N (%AAS)	3 (1.3)	13 (15.3)	8 (6.5)	10 (4.1)
Aortic rupture, N (%AAS)	0	0	26 (21.1)*	18 (7.5)
Alt. diagnoses, N (%tot)	802 (77.5)	291 (77.3%)	764 (86.1)	1607 (87)
Acute coronary syndrome, N (%AltD)	94 (12)	19 (6.5)	528 (69.1)	244 (13.2)
Stable angina, N (%AltD)	NR	35 (12.0%)	57 (7.5)	NR
Pulmonary embolism, N (%AltD)	13 (2)	42 (14.5)	29 (3.8)	30 (1.6)
Non complicated aortic aneurism, N (%AltD)	NR	0	NR	53 (2.9)
Syncope, N (%AltD)	66 (8)	0	NR	78 (4.2)
Pleuritis, N (%AltD)	NR	29 (10.0)	NR	57 (3.1)
Pericarditis, N (%AltD)	25 (3.1)	0	25 (3.3)	54 (2.9)
Limb/organ ischemia, N (%AltD)	12 (1.2)	0	1 (0.1)	2 (0.1)
Muscle-skeletal pain, N (%AltD)	302 (37.7)	166 (57.0)	NR	485 (26.2)
Gastrointestinal, N (%AltD)	73 (9.1)	0	5 (0.7)	191 (10.3)
Stroke, N (%AltD)	16 (2)	0	NR	15 (0.8)
Other, N (%AltD)	201 (19.4)	0	119 (15.6)	398 (21.5)

Supplementary Data File

AASs: acute aortic syndromes; Alt.: alternative; NR: not reported. *Includes rupture, impending-rupture and infectious aortic aneurysm. [§]Original data provided by the authors for the present analysis and not included in the original manuscript.

Supplementary Data File

Supplementary table 4. STARD 2015 checklist, modified from Bossuyt *et al*.

		Nazerian, 2014	Gorla, 2017	Kotani, 2017	Nazerian, 2018
1	Identification as a study of diagnostic accuracy using at least one measure of accuracy	Yes	Yes	Yes	Yes
2	Structured summary of study design, methods, results, and conclusions	Yes	Yes	Yes	Yes
3	Scientific and clinical background, including the intended use and clinical role of the index test	Yes	Yes	Yes	Yes
4	Study objectives and hypotheses	Yes	Yes	Yes	Yes
5	Whether is a prospective or retrospective study	Yes	Yes	Yes	Yes
6	Eligibility criteria	Yes	Yes	Yes	Yes
7	On what basis potentially eligible participants were identified	Yes	Yes	Yes	Yes
8	Where and when potentially eligible participants were identified	Yes	Yes	Yes	Yes
9	Whether participants formed a consecutive, random or convenience series	Yes	Unclear	Unclear	Yes
10a	Index test, in sufficient detail to allow replication	Yes	Yes	Yes	Yes
10b	Reference standard, in sufficient detail to allow replication	Yes	Unclear	Yes	Yes
11	Rationale for choosing the reference standard (if alternatives exist)	Yes	Unclear	Yes	Yes
12a	Definition of and rationale for test positivity cut-offs or result categories of the index test	Yes	Yes	Yes	Yes
12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard	Yes	Yes	Yes	Yes
13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	Yes	Yes	Yes	Yes

Supplementary Data File

13b	Whether clinical information and index test results were available to the assessors of the reference standard	Yes	Yes	Yes	Yes
14	Methods for estimating or comparing measures of diagnostic accuracy	Yes	Yes	Yes	Yes
15	How indeterminate index test or reference standard results were handled	Yes	Yes	Yes	Yes
16	How missing data on the index test and reference standard were handled	Yes	Unclear	Unclear	Yes
17	Any analyses of variability in diagnostic accuracy	Yes	Yes	Yes	Yes
18	Intended sample size and how it was determined	No	No	No	Yes
19	Flow of participants, using a diagram	Yes	Yes	Yes	Yes
20	Baseline demographic and clinical characteristics of participants	Yes	Yes	Yes	Yes
21a	Distribution of severity of disease in those with the target condition	Yes	Yes	Yes	Yes
21b	Distribution of alternative diagnoses in those without the target condition	Yes	Yes	Yes	Yes
22	Time interval and any clinical interventions between index test and reference standard	No	No	No	No
23	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Yes	Yes	Yes	Yes
24	Estimates of diagnostic accuracy and their precision	Yes	Yes	Yes	Yes
25	Any adverse events from performing the index test or the reference standard	No	No	No	No
26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	Yes	Yes	Yes	Yes
27	Implications for practice, including the intended use and clinical role of the index test	Yes	Yes	Yes	Yes
28	Registration number and name of registry	No	No	No	Yes
29	Where the full study protocol can be accessed	No	No	No	Yes
30	Sources of funding and other support; role of funders	Yes	Yes	Yes	Yes

Supplementary Data File

Supplementary table 5. Contingency table and meta-analysis of sensitivity and specificity values in the included studies, excluding patients with ADD-RS=0, for ADD-RS=1 and DD<500 ng/mL.

	TP	FP	FN	TN	Total	Sensitivity (95% CI)	Specificity (95% CI)
Nazerian 2014	212	349	2	150	713	99.1% (96.7% - 99.9%)	30.1% (26.1% - 34.3%)
Gorla 2017	82	48	2	55	187	97.6% (91.7% - 99.7%)	53.4% (43.3% - 63.3%)
Kotani 2017	111	219	1	139	470	99.1% (95.1% - 99.9%)	38.8% (33.7% - 44.1%)
Nazerian 2018	227	554	2	628	1411	99.1% (96.9% - 99.9%)	53.1% (50.2% - 56.0%)
Pooled estimate						98.7% (97.6% - 99.4%)	43.5% (31.2% - 56.3%)
I^2						0% (p = 0.74)	96.6% (p<0.001)

Heterogeneity was determined using the Higgins' I^2.

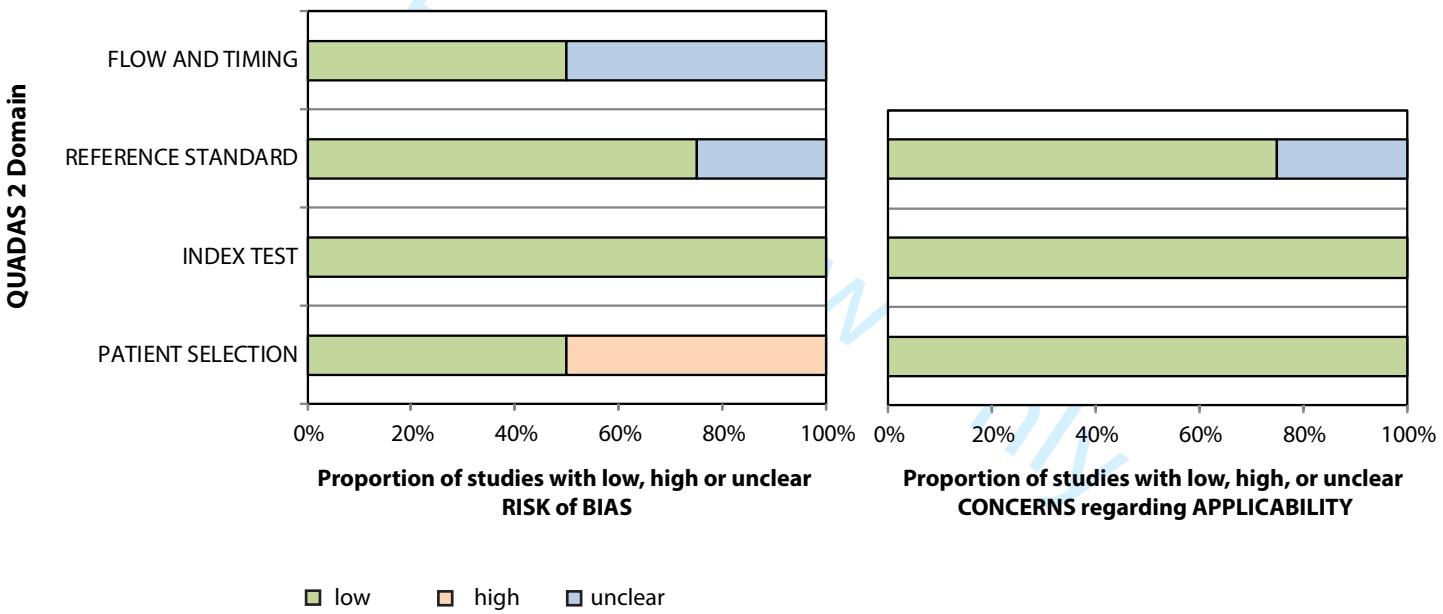
Supplementary Data File

Supplementary table 6. Contingency table and meta-analysis of sensitivity and specificity values in the included studies, excluding patients with ADD-RS=0, for ADD-RS=1 and DD<$DD_{age-adj}$.

	TP	FP	FN	TN	Total	Sensitivity (95% CI)	Specificity (95% CI)
Nazerian 2014	207	324	7	175	713	96.7% (93.4% - 98.7%)	35.1% (30.9% - 39.4%)
Golra 2017	81	42	3	61	187	96.4% (89.9% - 99.3%)	59.2% (49.1% - 68.8%)
Kotani 2017	109	180	3	178	470	97.3% (92.4% - 99.4%)	49.7% (44.4% - 55.0%)
Nazerian 2018	225	487	4	695	1411	98.3% (95.6% - 99.5%)	58.8% (55.9% - 61.6%)
Pooled estimate						97.1% (95.7% - 98.3%)	50.5% (38.1% - 62.8%)
I^2						0% (p=0.68)	96.4% (p<0.001)

Heterogeneity was determined using the Higgins' I^2.

Supplementary figure 1. Graphical display of QUADAS-2 results.



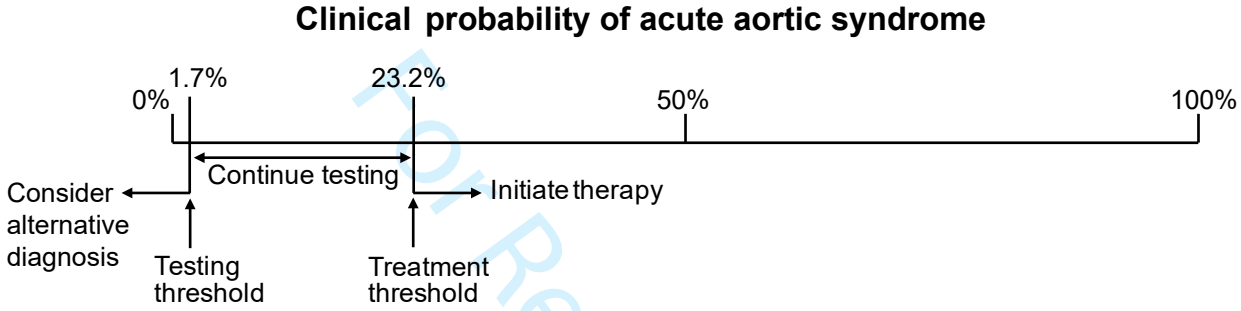
Supplementary figure 2. Test-treatment threshold analysis.

(a) ADD-RS =0 and D-dimer <500 ng/mL

$T_{testing\ threshold} = [(P_{pos/nd}) \times (R_{rx}) + R_t] \div [(P_{pos/nd} \times R_{rx}) + (P_{pos/d} \times B_{rx})] = 1.7\%$

$T_{treatment\ threshold} = [(P_{neg/nd}) \times (R_{rx}) - R_t] \div [(P_{neg/nd} \times R_{rx}) + (P_{neg/d} \times B_{rx})] = 23.2\%$

$P_{pos/nd}$ = probability of a positive result in patients without disease = 1 – specificity[a] = 1 – 0.19 = 0.81
$P_{neg/nd}$ = probability of a negative result in patients without disease = specificity[a] = 0.19
R_{rx} = risk of treatment in patients without disease = 0.010[b]
R_t = risk of diagnostic test = 0.0003[c]
$P_{pos/d}$ = probability of a positive result in patients with disease = sensitivity = 0.99
$P_{neg/d}$ = probability of a negative result in patients with disease = 1 – sensitivity = 1 – 0.99 = 0.01
B_{rx} = benefit of treatment in patients with disease = 0.50[d]

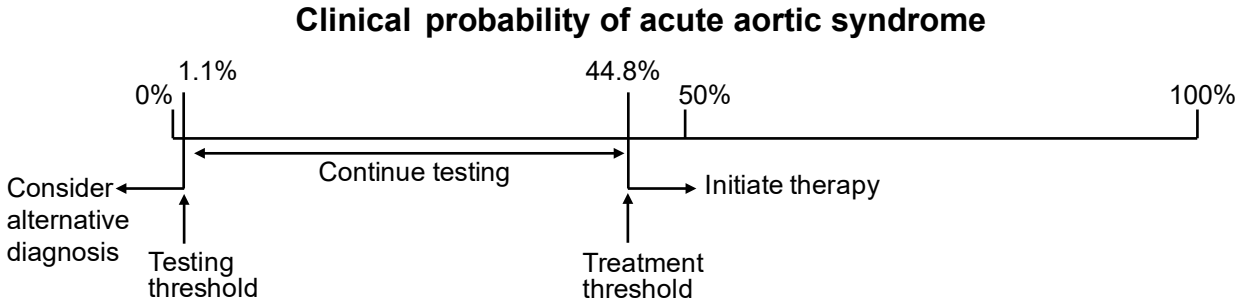


(b) ADD-RS ≤1 and D-dimer <500 ng/mL

$T_{testing\ threshold} = [(P_{pos/nd}) \times (R_{rx}) + R_t] \div [(P_{pos/nd} \times R_{rx}) + (P_{pos/d} \times B_{rx})] = 1.1\%$

$T_{treatment\ threshold} = [(P_{neg/nd}) \times (R_{rx}) - R_t] \div [(P_{neg/nd} \times R_{rx}) + (P_{neg/d} \times B_{rx})] = 44.8\%$

$P_{pos/nd}$ = probability of a positive result in patients without disease = 1 – specificity[a] = 1 – 0.46 = 0.54
$P_{neg/nd}$ = probability of a negative result in patients without disease = specificity[a] = 0.46
R_{rx} = risk of treatment in patients without disease = 0.010[b]
R_t = risk of diagnostic test = 0.0003[c]
$P_{pos/d}$ = probability of a positive result in patients with disease = sensitivity = 0.99
$P_{neg/d}$ = probability of a negative result in patients with disease = 1 – sensitivity = 1 – 0.99 = 0.01
B_{rx} = benefit of treatment in patients with disease = 0.50[d]



[a]pooled specificity
[b]based on Taylor RA, Iyer NS. A decision analysis to determine a testing threshold for computed tomographic angiography and D-dimer in the evaluation of aortic dissection. Am J Emerg Med 2013 Jul;31(7):1047-55.
[c]based on Cochran ST. Anaphylactoid reactions to radiocontrast media. Curr Allergy Asthma Rep. 2005 Jan;5(1):28-314
[d]estimated from mortality of treated and untreated type A (surgical treatment) and type B (medical treatment) aortic dissection (from Taylor RA et al.).