Comparison of clinical usefulness of serum Ca125 and CA19-9 in pancreatic adenocarcinoma diagnosis: meta-analysis and systematic review of literature.

Aleksander Skulimowski (aleksander.skulimowski@gmail.com)
Medical University of Lodz https://orcid.org/0000-0003-1968-7191
Adam Durczyński
Medical University of Lodz, Department of General and Transplant Surgery
Janusz Strzelczyk
Medical University of Lodz, Department of General and TRansplant Surgery
Piotr Hogendorf
Medical University of Lodz, Department of General and Transplant Surgery

Research article

Keywords: Pancreatic cancer, biomarkers, diagnosis, CA19-9, Ca125, meta-analysis

Posted Date: April 21st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-21433/v1

Version of Record: A version of this preprint was published at Biomarkers on January 31st, 2021. See the published version at https://doi.org/10.1080/1354750X.2021.1876770.

Abstract

Background: Pancreatic adenocarcinoma remains one of the most lethal cancers. This is caused by its late manifestation and lack of extensively available, highly accurate diagnostic tools. The only recommended biomarker CA19-9 proves to be not accurate enough to establish a certain diagnosis. Therefore a standardization of usefulness of other biomarkers is essential. Consequently, our aim was to assess the specificity and sensitivity of Ca125 in comparison with CA19-9 by means of meta-analysis. The systematic review of combined tests (CA19-9+Ca125) was also carried out.

Methods: We conducted a systematic search of Medline (via PubMed) and Ovid. After screening of abstracts and the assessment of full-texts, 9 studies (number of patients, n = 1599) were included in the meta-analysis. Hierarchical summary receiver under operator curve (hsROC) model was applied to estimate the diagnostic accuracy.

Results: CA19-9 sensitivity and specificity were 0,748 (95%CI 0,676-0,809) and 0,782 (95%CI 0,716-0,836) respectively. These values were estimated on 0,593 (95%CI 0,489-0,69) and 0,754 (95%CI 0,817-0,668) for Ca125. Regarding the heterogeneity of studies we found a strong threshold effect for Ca125 and moderate one for CA19-9.

Conclusions: Our meta-analysis did not prove the superiority of Ca125 in diagnosis of pancreatic adenocarcinoma. It should be nevertheless noted that analysed studies are encumbered with high heterogeneity mainly due to threshold effect. Moreover the sparsity of studies precludes accurate analysis of various factors' influence. Therefore further research into Ca125 is warranted to fully elucidate its usefulness. The review of proposed combined tests shows that although CA19-9+Ca125 models are characterized by higher sensitivity, their usefulness is hampered by inferior specificity to than that of CA19-9 alone.

Background

Pancreatic ductal adenocarcinoma (PDAC) belongs to the most lethal cancer. Despite its relatively low incidence, according to the latest cancer statistics it constitutes the 7th leading cause of cancer-related deaths worldwide. PDAC incidence is notably higher in the countries with high HDI (human development index) [1]. This epidemiological situation is caused essentially by lack of early specific symptoms and thus late diagnosis of PDAC. Secondly, the clinicians do not have at their disposal any readily available diagnostic test to ascertain or exclude PDAC diagnosis with high probability. In the current European oncological guidelines of European Society of Medical Oncology (ESMO) CA19-9 still remains as the sole recommended serum biomarker [2]. Nevertheless, its shortcomings are well-known. (i) Around 10% of Caucasian population are so called "Lewis-antigen non-expressors" what leads in turn to no expression of CA19-9 [3]. (ii) Surge of CA19-9 levels is fairly often seen in plethora of other diseases [4]. (iii) Early stage of PDAC is often seen without increase of CA19-9. The conducted meta-analysis concluded that CA19-9 sensitivity and specificity are around 80% [5]. Both false positive and false negative diagnosis have

serious ramifications. While false negative diagnosis leads, obviously, to the delay of oncological treatment, the false positive result, in the clinical scenario of pancreatic mass's presence of benign etiology, can be similarly detrimental. In such a scenario, patients with the misdiagnosed PDAC receive unnecessarily extensive surgery such as pancreatoduodenectomy or pancreatomy. According to published studies the rate of pancreatoduodenectomy due to misdiagnosed PDAC amount to 5-12% of cases [6-9]. Moreover, routinely used imaging studies also lack exceptionally highly specific and sensitive in differentiating between pancreatic mass due to e.g. chronic pancreatits and PDAC [10-11]. Ca125, similarly as CA19-9, belongs to the high-mass glycoproteins. Its clinical usefulness was firstly described in diagnosis of ovarian cancer [12]. Up-to-date Ca125 and HE4 constitute as two independent factors in ROMA (Risk of Ovarian Malignancy Algorithm) test, which is used for the calculation of the probability of ovarian cancer presence [13]. Nevertheless, there is a growing number of evidence that Ca125 is also up-regulated in the development of PDAC. In our previous retrospective study we also found that Ca125 with the optimal cut-off point has the diagnostic accuracy matching that of CA19-9 [14]. Therefore we aimed to conduct the meta-analysis and the systematic review of literature to assess the Ca125 performance against CA19-9.

Methods

Search strategy

The performed meta-analysis and systematic review were in line with the the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [15].We performed the search of the following databases: MEDLINE (via PubMed), EMBASE (via Ovid). The last search was performed on 25th of February 2020. We used the following search construct: pancreatic cancer OR pancreatic tumor OR pancreatic adenocarcinoma OR pancreatic lesion AND diagnosis AND ca125 AND CA19-9. The number of identified potentially eligible studies is shown in the PRISMA Flow Diagram (Figure 1).

Eligibility criteria

We established a priori the following eligibility criteria:

i. Case-control or diagnostic cohorts
ii. Studies published in English or German
iii. Study population of at least 60 participants
iv. Data included in the paper enabling to create 2x2 diagnostic table
v. Histopathological examination as a gold standard

Data extraction and study inclusion

Two authors (A.S. and A.D.) independently screened the

records retrieved from the search. Selected records were further

screened for eligibility in full text independently by the same investigators. Discrepancies at each stage of selection were arbitrated by a third reviewer (P.H.) and resolved by consensus.

Assessment of methodological quality

In order to assess the quality of each study the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) were applied [16]. Two authors (A.S. and A.D.) independently filled out the assessment form for each included study. Discrepancies in the assessment were resolved by the third author (P.H.).

Data preparation

After extracting the full-text of each included study, we built 2x2 diagnostic tables with false positive, negative (FP and FN) and true positive and negative (TP and TN) rates. Moreover, for each study, the information about geographical origin of the population, whether Ca125 optimal cut-off point was calculated, method of biomarkers' detection, were coded as categorical variables for the further exploration by meta-regression. The data is shown in Table 2.

Publication bias

To statistically assess the possible publication bias, the diagnostic odds ratios (DOR) for each study were calculated. Then, natural logarithm transformation was performed (lnDOR), as well as the calculation of standard error (SE). Finally funnel plots were constructed by plotting Precision ($1/SE^{-1}$) against lnDOR. Trim and fill test was applied to evaluate the asymmetry of resulting plot.

Meta-Analysis Methodology

Currently univariate analysis methods are not recommended in data synthesis of diagnostic tests' studies [17, 18]. Although there is no clear consensus regarding the optimal analysis methods, it is widely agreed that hierarchical summary receiver under operator curve (hsROC) are the best suited for this purpose [19]. Therefore we decided to use the model of hsROC, firstly described by Reitsma et al. [20]. Further this model was applied to assess pooled diagnostic accuracy for each biomarker.

In terms of the exploration of studies' heterogeneity and subgroup analysis the parametric transformation described by Dobbler et al. [21] was applied for the bivariate meta-regression with maximal likelihood estimation.

Univariate approach was used for the graphical summary of the sensitivity and specificity of each study as a forest plot (random effects model was applied), however the pooled specificity and sensitivity were generally not taken into account when comparing the biomarkers' performance.

All calculations were performed using the R programming language [22] and two packages dedicated for calculations on meta-data, namely mada [23] and metafor [24] were used.

Results

Studies' quality assessment and publication bias.

Simultaneously to the primary data extraction the risk of studies' bias was evaluated as proposed by QUADS-2 tool (Table 1.). In case of 5 studies (55,56%) the risk of bias in the patients selection domain was assessed as high, due to case-control design (4 studies) and exclusion of patients without "confident clinical diagnosis". Regarding the applicability in this domain we found control groups' composition in 2 studies (22,23%) as bias-prone, since in two cases they comprised, to some extent, of acute pancreatitis and extra-pancreatic cases [29,31]. Normally acute pancreatitis is clinically easily distinguishable from PDAC, thus its enrollment to the control group doesn't seem to be fully justified. In the other case the bias is attributed to the inclusion of the pancreatic neuroendocrine tumors in the cancer group [26]. Lastly, the study by Gu et al. [25] included exclusively PDAC patients undergoing chemotherapy. It was not stated clearly whether the sample taken for the diagnostic purposes was obtained before the commencement of chemotherapy in every case, thus we assessed the risk as unclear.

In the index test domain two studies have unclear bias risk, as the authors do not state clearly whether the test interpretation was "blinded" to the results of reference test. Additionally in one case the authors used two distinct assays to measure Ca125 levels [25-27]. In one case we assess the applicability of index test as low, since the authors used two different cut-off points for Ca125 [28].

In the remaining domains and applicability concerns we evaluated the bias risk as low.

The funnel plots are shown in Supplementary Figure 1 and 2. The performed trim and fill method excluded the plot asymmetry for both biomarkers (p = 0,14 for CA19-9 and p = 0,11 for Ca125).

Meta-analysis

We identified 230 potential articles through various literature databases. After removing duplicates and irrelevant ones, 22 studies remained. These were screened by abstract and/or full-text for the eligibility. After reviewing them basing on our criteria, finally 9 studies were included to the meta-analysis [14, 25-32]. The detailed flow diagram is depicted in Figure 1.

The conducted meta-analysis included 4 European studies, 1 from the United States and 4 Asian studies. 5 studies were designed as cohort studies, while 4 of them had case-control design (Table 2). They included overall 1599 patients, of whom 975 had PDAC (61%), while the control group consisted of 624 patients (39%). 261 of them had chronic pancreatitis, 102 other benign pancreatic diseases/other benign diseases [25, 31], 77 acute pancreatitis, 50 cholelithiasis, 41 pancreatic cyst, 23 cholangiocarcinoma, 19 pancreatic pseudocyst, 10 pancreatic cystic neoplasm and 1 patient was diagnosed with pancreatic arteriovenous malformation. Furthermore one study enrolled 40 healthy patients [25].

The summary forest plots are shown in Figure 2. As depicted, the studies vary significantly regarding reported sensitivity and specificity for both CA19-9 and Ca125.

Additionally we calculated diagnostic odds ratio, positive and negative likelihood ratios for all included studies. The results are presented in the Table 3.

We then calculated hierarchical summary ROC for both biomarkers. The curves are shown in Figure 3.

The point estimate for CA19-9 has the following parameters:

Sensitivity: 0,748 [95%CI: 0,676-0,809]

Specificity: 0,782 [95%CI: 0,716-0,836]

Area Under Curve (AUC) was estimated for 0,832.

Using the calculated hsROC, we applied it to further calculate the mean DOR, PLR and NLR.

Diagnostic Odds Ratio: 10,9 (7,56-15,1)

Positive Likelihood Ratio: 3,46 (2,72-4,4)

Negative Likelihood Ratio: 0,324 (0,252-0,403)

These parameters have the following values for Ca125:

Sensitivity: 0,593 [95%CI: 0,489-0,69]

Specificity: 0,754 [95%CI: 0,678-0,817]

AUC: 0,739

Diagnostic Odds Ratio: 4,52 (3,41-5,88)

Positive Likelihood Ratio: 2,42 (2,01-2,92)

Negative Likelihood Ratio: 0,541 (0,441-0,641)

As shown in the curve comparison (Figure 3), the points of estimate are well separated, with only a few studies overlapping, suggesting that CA19-9 has indeed significantly better performance over Ca125. Nevertheless we aimed to elucidate the heterogeneity influence on the pooled diagnostic accuracy.

As suggested by others authors, Spearman correlation between sensitivity and false positive rate (fpr) was calculated. The Spearman rho was 0,545 and 0,764 for CA19-9 and Ca125, respectively, indicating a possible significant threshold effect for Ca125 (rho >= 0,7).

Heterogeneity analysis

To further explore the studies' heterogeneity, we performed a meta-regression. We chose a priori the following factors as a possible sources of heterogeneity:

i. Calculated cut-off point for Ca125 vs standard cut-off point
ii. Study location (Asia vs. Europe/USA)
iii. Study type (cohort vs. case-control studies)
iv. Publication year (before vs. after 2010)
v. Method of biomarker assessment
vi. PDAC prevalence in the study population

As shown in Table 2, all the studies published before 2010 used a type of radioimmunoassay for the biomarkers' assessment, thus studies' split regarding points IV and V is same.

We did not found any statistically significant impact of study location, type, publication year (e.g. method of biomarker assessment) on sensitivity or specificity of Ca125 (Supplementary Table 1.). However, the built meta-regression model showed that studies with calculated cut-off point and higher PDAC prevalence (>60%) tend to report higher sensitivity for Ca125 (p=0,021 and 0,04 respectively). To further assess the significance of these difference the likelihood-ratio test was performed, that concluded the differences between bivariate models (general parametric model vs. parametric model with a covariate) as insignificant (p= 0,153 and p = 0,2 respectively). Similarly, in the univariate subgroup analysis the calculated differences were insignificant. The pooled sensitivity and specificity for the studies estimating cut-off point value for Ca125 (n=3) were 0,696 (0,573-0,796) and 0,676 (0,53-0,794) respectively. For the studies without optimal cut-off point estimation, these values were 0,539 (0,432-0,642) and 0,784 (0,721-0,836) (p = 0,055 and 0,056 respectively).

Interestingly, the meta-regression for CA19-9 revealed that studies with a calculated cut-off point for Ca125 reported lower sensitivity for CA19-9 (p < 0,0001), while studies conducted in Europe/USA had significantly lower sensitivity and significantly higher specificity than the Asian ones (p = 0,032 and p = 0,038 respectively). Finally, the older studies (before 2010) were characterized by higher sensitivity (p = 0,016) (Supplementary Table 2.). However, the conducted likelihood-ratio test did not confirm the significance of the observed differences (p = 0,16, p = 0,2 and p = 0,075 respectively).

Systematic review of combined diagnostic tests

The designed tests are summarized in the Table 4. Apart from the study by Wang et al., all the reviewed articles proposed a combination test of Ca125 with the other measured biomarkers. Four older papers examined simple AND/OR formulae, that took into account CA19-9 and Ca125 levels. While the application of AND formula caused a significant increase in specificity of test with concomitant decrease of sensitivity, OR formula had an inverse impact on test's parameters. Though maximalization of one parameter at cost of another might seem promising, in all cases, apart from the model from study by Sakamoto et al. (using AND formula), the accompanying decrease was greater than resulting increase, so that the proposed combinations didn't outperformed the diagnostic accuracy of CA19-9. On the other hand three more recent studies used a logistic regression model. All the designed test succeeded in improving sensitivity over CA19-9. While the test constructed by Chan et al. managed to outperform

CA19-9 sensitivity without any “loss” on specificity, both combination models devised in our department does it at the cost of significantly lower specificity.

The test reported by Gu et al. stands somewhat apart from the other combinations, as the reported joint detection of CA19-9, Ca125, CEA and CA242 should lead to increase of both sensitivity and specificity. Unfortunately, the authors did not provide any information about the mathematical rationale behind their test.

Conclusions

The conducted meta-analysis did not find the superiority of Ca125 over CA19-9 in the diagnosis of PDAC. It should be however noted that due to the sparsity of studies comparing the both biomarkers, high heterogeneity and different control groups, the results should be taken with a certain amount of skepticism.

From the clinical point of view, one of the most important factors contributing to the differences in the estimation of Ca125 accuracy is the study’s design. Part of the included studies, such as ours, dealt with the diagnosis of etiology of the encountered pancreatic mass. Others, especially case control study, however, took more broad approach to the problem by comparing the biomarkers’ levels between PDAC and benign disease, often of extra-pancreatic genesis. In our opinion the core question is rather whether the etiology of pancreatic mass can be ascertained by measurement of serum biomarkers, as benign diseases require obviously less invasive treatment methods and, as already stated, the misdiagnosis can have grave consequences for the patients.

Further, it should be noted that most of the studies included have fairly moderate study population and in two cases there is an underrepresentation of PDAC cases.

Apart from the significant threshold-effect for Ca125, in the conducted meta-regression we did not find any other significant factors contributing to the heterogeneity. Nevertheless, it should be noted that there is some evidence showing that the disease’s prevalence in the study population has an impact on sensitivity and specificity of the conducted diagnostic test [33,34]. As for the optimal cut-off calculation, the methods like Younden’s index enable to find an optimal one to maximize the diagnostic accuracy. Unfortunately, most of the included studies used merely the cut-off point suggested by the test’s manufacturer protocol. It would be interesting to compare the diagnostics accuracy resulting from the application of optimal cut-off point with the parameters calculated for recommended cut-off point.

In the recent years Ca125 emerged as a promising target in pancreatic cancer research. Of note, there is a growing number of evidence that MUC-16, from which Ca125 originates, can have a vital role in PDAC development [35, 36]. Due to that, the MUC16-targetting in PDAC immunotherapy might be promising [37, 38].

Regarding the analysis of heterogeneity in CA19-9, the preliminary observed difference between reported sensitivity and specificity in European/USA and Asian studies could be attributed to different prevalence of Lewis-antigen non-secretors in the populations. Our study form 2018, where the CA19-9 sensitivity of 52,38% was reported, may account for the observed difference between studies with the calculated cut-off point for Ca125 (n = 3) and those without (n = 6). These differences did not prove, however, to be significant when comparing the bivariate models.

Furthermore, the impact of low number for studies enrolled to the meta-analysis should be acknowledged. Firstly, it definitely influences the conducted publication bias analysis, as a rule of thumb states that 10 or more studies are required to perform an accurate one [39]. Secondly, the sparsity of studies can lead to difficult hsROC model fit and result in the unreliable estimation of parameters [40].

The point estimate in the hsROC model for CA19-9 has sensitivity of around 75% and specificity of 78%, what is quite similar to the results from the previous meta-analysis [5]. While these numbers indicate fairly good diagnostic accuracy of CA19-9, they are definitely too low to accept it as a standard for PDAC diagnosis. Thus, further research into pancreatic cancer biomarkers is crucial to the improvement of current epidemiology.

While CA19-9 and Ca125 are normally tested in the scenario of “immediate” diagnosis of pancreatic cancer, O’Brien et al. [41] analyzed serum levels of CA19-9 and Ca125 of 458 post-menopausal women. 154 of them were subsequently diagnosed with PDAC, the rest served as matched non-cancer control. The authors proved that a model of CA19-9 > 37 IU/mL OR Ca125 >30 IU/mL has a sensitivity of 95,2% and specificity of 57,1% in “diagnosing” PDAC 0-1 year (average time: 6 months) before the initial diagnosis was made.

It should be also noted that usefulness of both CA19-9 and Ca125 goes beyond the diagnosis of PDAC. A growing number of studies shows that monitoring of CA19-9 and Ca125 can serve as a prognostic factor of survival or as an indicator of recurrence. The aforementioned study by O’Brien et al. showed that patients with CA19-9 > 40 IU/mL had median survival time from sample collection of 14.5 versus 36.0 months for non-elevated group. Ca125 > 25 IU/mL was correlated with a median survival time of 14 moth versus 35 months. In other recent study preoperative Ca125 >= 18,4 IU/mL was associated with poorer surgical outcomes [42]. There is also an ample evidence that both biomarkers can serve as predicators of chemotherapy response and recurrence of PDAC [43-45].

To conclude, the gathered evidence is rather insufficient to undoubtedly state that Ca125 is significantly inferior to CA19-9 in terms of diagnostic accuracy. The most important problem here is the studies’ sparsity, as this can result in the suboptimal fit of hsROC model and lead to biased conclusions.

Nevertheless, since the hsROC curves are only minimally overlapping, the trend towards CA19-9 superiority, especially in case of higher sensitivity, should be appreciated.

In order to fully validate the usefulness of Ca125 in the diagnosis of PDAC the bigger, well-designed studies are paramount.

The review of combined tests shows that a fairly simplistic mathematical model like logistic regression applied to a CA19-9/Ca125-based biomarker panel can significantly increase the diagnostic accuracy. While the results of systematic review are insufficient to state whether a mere combination of CA19-9 and Ca125 would be enough to significantly increase the accuracy, a theoretical panel based either on one or both of them could prove to be extremely valuable due to its simplicity and cost-effectiveness.

Abbreviations

AUC- Area under curve

DOR- Diagnostic odds ratio

CEA- Carcinoembryonic antigen

ESMO- European Society of Medical Oncology

FN- False negative result

FP- False positive result

Fpr- false positive rate

HDI- human development index

lnDOR- logarithm transformation of diagnostic odds ratio value

hsROC- hierarchical summary receiver under operator curve

NLR- Negative likelihood ratio

PLR- Positive likelihood ratio

PDAC- Pancreatic ductal adenocarcinoma

PRIMSA- Preferred Reporting Items for Systematic Reviews and Meta-Analyses

ROMA- Risk of Ovarian Malignancy Algorithm

SE- Standard error

TN- True negative result

TP- True positive result

QUADS- Quality Assessment of Diagnostic Accuracy Studies

Declarations

Ethics approval and consent to participate

Not applicable (Ethics approval and consent to participate forms available for each enrolled study)

Consent for publication

Not applicable

Availability of data and materials

All data generated or analysed during this study available from the corresponding author (Aleksander Skulimowski) on personal request

Competing interests

The authors declare that they have no competing interests

Acknowledgements

Not applicable.

Funding

None

Authors' contributions

AS and AD acted as the principal investigators during the whole workflow of the meta-analysis (publications' search, inclusion into meta-analysis, evaluation of studies' quality). PH acted as the third investigator (arbitration by discrepancies). AS performed the calculations in R-language environment. AS and JS wrote the manuscript. All the authors reviewed the manuscript and accepted it for the publication.

References

1. Rawla P, Sunkara T, Gaduputi V. Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. World J Oncol. 2019 Feb;10(1):10-27.
2. Ducreux M, Cuhna AS, Caramella C et al. Cancer of the pancreas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2015 Sep;26 Suppl 5:v56-68.
3. Scarà S, Bottoni P, Scatena R CA 19-9: Biochemical and Clinical Aspects. Adv Exp Med Biol. 2015;867:247-60.

4. Goh SK, Gold G, Christophi C, Muralidharan V. Serum carbohydrate antigen 19-9 in pancreatic adenocarcinoma: a mini review for surgeons. ANZ J Surg. 2017 Dec;87(12):987-992.
5. Huang Z, Liu F. Diagnostic value of serum carbohydrate antigen 19-9 in pancreatic cancer: a meta-analysis. Tumour Biol. 2014 Aug;35(8):7459-65.
6. Gomes RM, Bal M, Patkar S, Goel M, Shrikhande SV. Unexpected benign histopathology after pancreatoduodenectomy for presumed malignancy: accepting the inevitable. Langenbecks Arch Surg. 2016 Mar;401(2):169-79.
7. Gerritsen A, Molenaar IQ, Bollen TL, et al. Preoperative characteristics of patients with presumed pancreatic cancer but ultimately benign disease: a multicenter series of 344 pancreatoduodenectomies. Ann Surg Oncol. 2014 Nov;21(12):3999-4006.
8. van Heerde MJ, Biermann K, Zondervan PE et al. Prevalence of autoimmune pancreatitis and other benign disorders in pancreatoduodenectomy for presumed malignancy of the pancreatic head. Dig Dis Sci. 2012 Sep;57(9):2458-65.
9. Kennedy T, Preczewski L, Stocker SJ et al. Incidence of benign inflammatory disease in patients undergoing Whipple procedure for clinically suspected carcinoma: a single-institution experience. Am J Surg. 2006 Mar;191(3):437-41.
10. Gerritsen A, Bollen TL, Nio Y et al. Diagnostic value of a pancreatic mass on computed tomography in patients undergoing pancreatoduodenectomy for presumed pancreatic cancer. Surgery. 2015 Jul;158(1):173-82.
11. Toft J, Hadden WJ, Laurence JM et al. Imaging modalities in the diagnosis of pancreatic adenocarcinoma: A systematic review and meta-analysis of sensitivity, specificity and diagnostic accuracy. Eur J Radiol. 2017 Jul;92:17-23. Doi: 10.1016/j.ejrad.2017.04.009.
12. Bottoni P, Scatena R. The Role of CA 125 as Tumor Marker: Biochemical and Clinical Aspects. Adv Exp Med Biol. 2015;867:229-44.
13. Dochez V, Caillon H, Vaucel E et al. Biomarkers and algorithms for diagnosis of ovarian cancer: CA125, HE4, RMI and ROMA, a review. J Ovarian Res. 2019 Mar 27;12(1):28.
14. Hogendorf P, Skulimowski A, Durczyński A et al. A Panel of CA19-9, Ca125, and Ca15-3 as the Enhanced Test for the Differential Diagnosis of the Pancreatic Lesion. Dis Markers. 2017;2017:8629712.
15. Moher D1, Liberati A, Tetzlaff J Altman DG and PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6(7): e1000097
16. 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 Oct 18;155(8):529-36.
17. Leeflang MM. Systematic reviews and meta-analyses of diagnostic test accuracy. Clin Microbiol Infect. 2014 Feb;20(2):105-13.
18. Lee J, Kim KW, Choi SH et al. Systematic Review and Meta-Analysis of Studies Evaluating Diagnostic Test Accuracy: A Practical Review for Clinical Researchers-Part II. Statistical Methods of Meta-Analysis. Korean J Radiol. 2015 Nov-Dec;16(6):1188-96.

19. Dinnes J, Mallett S, Hopewell S et al. The Moses-Littenberg meta-analytical method generates systematic differences in test accuracy compared to hierarchical meta-analytical models. J Clin Epidemiol. 2016 Dec;80:77-87.
20. Reitsma JB, Glas AS, Rutjes AW et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol. 2005 Oct;58(10):982-90.
21. Doebler P, Holling H, B□ohning D (2012). \A Mixed Model Approach to Meta-Analysis of Diagnostic Studies With Binary Test Outcome." Psychological Methods.
22. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
23. Philipp Doebler (2019). Mada: Meta-Analysis of Diagnostic Accuracy. R package version 0.5.9. https://CRAN.R-project.org/package=mada
24. Balduzzi S, Rücker G, Schwarzer G (2019), How to perform a meta-analysis with R: a practical tutorial, Evidence-Based Mental Health.
25. Gu YL, Lan C, Pei H et al. Applicative Value of Serum CA19-9, CEA, CA125 and CA242 in Diagnosis and Prognosis for Patients with Pancreatic Cancer Treated by Concurrent Chemoradiotherapy. Asian Pac J Cancer Prev. 2015;16(15):6569-73.
26. Wang Z, Tian YP. Clinical value of serum tumor markers CA19-9, CA125 and CA72-4 in the diagnosis of pancreatic carcinoma. Mol Clin Oncol. 2014 Mar;2(2):265-268
27. Duraker N, Hot S, Polat Y et al. CEA, CA 19-9, and CA 125 in the differential diagnosis of benign and malignant pancreatic diseases with or without jaundice. J Surg Oncol. 2007 Feb 1;95(2):142-7.
28. Sakamoto K, Haga Y, Yoshimura R et al. Comparative effectiveness of the tumour diagnostics, CA 19-9, CA 125 and carcinoembryonic antigen in patients with diseases of the digestive system. Gut. 1987 Mar;28(3):323-9.
29. Haglund C. Tumour marker antigen CA125 in pancreatic cancer: a comparison with CA19-9 and CEA. Br J Cancer. 1986 Dec;54(6):897-901.
30. Cwik G, Wallner G, Skoczylas T et al. Cancer antigens 19-9 and 125 in the differential diagnosis of pancreatic mass lesions. Arch Surg. 2006 Oct;141(10):968-73;
31. Chan A, Prassas I, Dimitromanolakis A et al. Validation of biomarkers that complement CA19.9 in detecting early pancreatic cancer. Clin Cancer Res. 2014 Nov 15;20(22):5787-95.
32. Hogendorf P, Durczyński A, Skulimowski A et al. Growth differentiation factor (GDF-15) concentration combined with Ca125 levels in serum is superior to commonly used cancer biomarkers in differentiation of pancreatic mass. Cancer Biomark. 2018 Feb 14;21(3):505-511.
33. Leeflang MM, Rutjes AW, Reitsma JB et al. Variation of a test's sensitivity and specificity with disease prevalence. CMAJ. 2013 Aug 6;185(11):E537-44. doi: 10.1503/cmaj.121286. Epub 2013 Jun 24.
34. Leeflang MM, Bossuyt PM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. J Clin Epidemiol. 2009 Jan;62(1):5-12.

35. Das S, Rachagani S, Torres-Gonzalez MP et al. Carboxyl-terminal domain of MUC16 imparts tumorigenic and metastatic functions through nuclear translocation of JAK2 to pancreatic cancer cells. Oncotarget. 2015 Mar 20;6(8):5772-87.
36. Liu Q, Cheng Z, Luo L et al. C-terminus of MUC16 activates Wnt signaling pathway through its interaction with β-catenin to promote tumorigenesis and metastasis. Oncotarget. 2016 Jun 14;7(24):36800-36813
37. Garg G, Gibbs J, Belt B et al. Novel treatment option for MUC16-positive malignancies with the targeted TRAIL-based fusion protein Meso-TR3. BMC Cancer. 2014 Jan 21;14:35.
38. Aithal A, Rauth S, Kshirsagar P et al. MUC16 as a novel target for cancer therapy. Expert Opin Ther Targets. 2018 Aug;22(8):675-686.
39. Dalton JE, Bolen SD, Mascha EJ. Publication Bias: The Elephant in the Review. Anesth Analg. 2016 Oct;123(4):812-3.
40. Takwoingi Y, Guo B, Riley RD, Deeks J. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. Stat Methods Med Res. 2017 Aug;26(4):1896-1911.
41. O'Brien DP, Sandanayake NS, Jenkinson C et al. Serum CA19-9 is significantly upregulated up to 2 years before diagnosis with pancreatic cancer: implications for early disease detection. Clin Cancer Res. 2015 Feb 1;21(3):622-31.
42. Liu L, Xu HX, Wang WQ et al. Serum CA125 is a novel predictive marker for pancreatic cancer metastasis and correlates with the metastasis-associated burden. Oncotarget. 2016 Feb 2;7(5):5943-56.
43. Xu HX, Li S, Wu CT et al. Postoperative serum CA19-9, CEA and CA125 predicts the response to adjuvant chemoradiotherapy following radical resection in pancreatic adenocarcinoma. Pancreatology. 2018 Sep;18(6):671-677.
44. Nishio K., Kimura K., Amano R., et al. Preoperative predictors for early recurrence of resectable pancreatic cancer. World J. Surg. Oncol.: 2017; 15(1): 16
45. Xu HX, Liu L, Xiang JF et al. Postoperative serum CEA and CA125 levels are supplementary to perioperative CA19-9 levels in predicting operative outcomes of pancreatic ductal adenocarcinoma. Surgery. 2017 Feb;161(2):373-384.

Tables

Study	RISK OF BIAS				APPLICABILITY CONCERNS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Halgund et al.	L	J	J	J	L	J	J
Sakamoto et al.	L	J	J	J	J	L	J
Cwik et al.	J	J	J	J	J	J	J
Duraker et al.	L	L	?	J	J	L	J
Wang et al.	J	?	?	J	L	J	J
Chan et al.	L	J	J	J	L	J	J
Gu et al.	L	?	?	J	?	J	J
Hogendorf et al. 2017	J	J	J	J	J	J	J
Hogendorf et al. 2018	J	J	J	J	J	J	J

JLow Risk LHigh Risk ? Unclear Risk

Table 1. The primary studies' assessment by QUADS-2 tool.

Study	Year of publication	Type of study	Country of origin	Biomarker detection method	Ca125 optimal cut-off point calculated? (Y/N)	PDAC prevalence in the study group [%]	Total study population	TP	FP	FN	TN
Halgund et al.	1986	Case control	Finland	RIA	N	53,37	178	74	18	21	65
Sakamoto et al.	1987	Case-control	Japan	RIA	N	35,71	84	26	16	4	38
Cwik et al.	2006	Cohort	Poland	RIA	N	66,36	110	59	4	14	33
Duraker et al.	2007	Cohort	Turkey	RIA	N	67,96	181	100	14	23	44
Wang et al.	2013	Cohort	China	ECLIA	N	51,72	145	55	10	20	60
Chan et al.	2014	Case-control	USA	ELISA	Y	65,79	380	182	24	68	106
Gu et al.	2015	Case-control	China	ECLIA	N	39,39	132	43	33	9	47
Hogendorf et al.	2017	Cohort	Poland	ELFA	Y	71,78	326	138	19	96	73
Hogendorf et al.	2018	Cohort	Poland	ELFA	Y	66,67	63	24	2	18	19

Table 2. The overview of the included studies

RIA – Radioimmuno Assay, ECLIA – Electro-Chemiluminescence Immunoassay ELISA – Enzyme-Linked Immunosorbent Assay, ELFA - Enzyme-Linked Fluorescent Assay

Study	Biomarker	Sensitivity	Specificity	DOR	Positive LR	Negative LR
Halgund et al.	**Ca125**	0,45 (0,36-0,55)	0,76 (0,66-0,84)	2,6 (1,37-4,97)	1,88 (1,2-2,92)	0,72 (0,58-0,9)
	CA19-9	**0,78** (0,69-0,85)	**0,78** (0,68-0,86)	12,73 (6,24-25,94)	3,59 (2,35-5,48)	0,28 (0,2-0,42)
Sakamoto et al.	**Ca125**	0,63 (0,46-0,78)	**0,78** (0,65-0,87)	6,05 (2,27-16,13)	2,85 (1,61-5,03)	0,47 (0,29-0,78)
	CA19-9	**0,87** (0,7-0,95)	0,7 (0,57-0,81)	15,44 (4,63-51,45)	2,93 (1,9-4,52)	0,19 (0,08-0,48)
Cwik et al.	**Ca125**	0,61 (0,49-0,71)	0,83 (0,68-0,92)	7,76 (2,87-20,94)	3,65 (1,72-7,75)	0,47 (0,34-0,65)
	CA19-9	**0,81** (0,7-0,88)	**0,89** (0,75-0,96)	34,77 (10,58-114,29)	7,48 (2,94-19)	0,22 (0,13-0,35)
Duraker et al.	**Ca125**	0,57 (0,48-0,65)	**0,78** (0,65-0,86)	4,57 (2,24 -9,33)	2,54 (1,54-4,2)	0,56 (0,43-0,71)
	CA19-9	**0,81** (0,74-0,87)	0,76 (0,63-0,85)	13,67 (6,44-29,02)	3,37 (2,12-5,36)	0,25 (0,17-0,37)
Wang et al.	**Ca125**	0,3 (0,21-0,42)	**0,89** (0,79-0,94)	3,43 (1,42-8,3)	2,68 (1,29-5,6)	0,78 (0,66-0,93)
	CA19-9	**0,73** (0,62-0,82)	0,86 (0,76-0,92)	16,5 (7,1-38,32)	5,13 (2,85-9,26)	0,31 (0,21-0,46)
Chan et al.	**Ca125**	0,7 (0,64-0,75)	0,75 (0,67-0,82)	7,15 (4,41-11,57)	2,84 (2,08-3,88)	0,4 (0,32-0,49)
	CA19-9	**0,73** (0,67-0,78)	**0,82** (0,74-0,87)	11,82 (7-22)	3,94 (2,73-5,7)	0,33 (0,27-0,42)
Gu et al.	**Ca125**	0,69 (0,56-0,8)	**0,68** (0,57-0,77)	4,67 (2,2-9,91)	2,13 (1,48-3,07)	0,46 (0,3-0,7)
	CA19-9	**0,83** (0,7-0,91)	0,59 (0,48-0,69)	6,81 (2,92-15,84)	2,01 (1,5-2,68)	0,3 (0,16-0,55)
Hogendorf et al. (2017)	**Ca125**	**0,79** (0,73-0,84)	0,52 (0,42-0,62)	4,12 (2,46-6,9)	1,65 (1,32-2,07)	0,4 (0,29-0,55)
	CA19-9	0,59 (0,53-0,65)	**0,79** (0,7-0,86)	5,52 (3,13-9,75)	2,86 (1,89-4,32)	0,52 (0,43-0,62)
Hogendorf et al. (2018)	**Ca125**	0,52 (0,38-0,67)	0,76 (0,55-0,89)	3,52 (1,09-11,37)	2,2 (0,97-4,98)	0,63 (0,42-0,93)
	CA19-9	**0,57** (0,42-0,71)	**0,9** (0,71-0,97)	12,67 (2,61-61,5)	6 (1,57-23)	0,47 (0,33-0,7)

Table 3. The estimated diagnostic parameters. 95% confidence intervals are given in the brackets.

Study	Combined biomarkers	Test formula	Sensitivity	Specificity
Halgund et al.	CA19-9 and Ca125	AND/OR Formulae	41% (-37%)/84% (6%)	93% (15%)/60% (-18%)
Sakamoto et al.	CA19-9 and Ca125	AND Formula	96,67% (9,67%)	64,81% (-5,19%)
Cwik et al.	CA19-9 and Ca125	OR Formula	87,8% (7%)	77,8% (-11,3%)
Duraker et al.	CA19-9 and Ca125	AND/OR Formulae	46,3% (-35%)/91,9% (10,6%)	93,1%(17,2%)/60,3% (-15,6%)
Wang et al. $^{\sigma}$	CA19-9 and CA72-4	Logistic regression: 1.496 * 0.004xCA19-9 * 0.207xCA72-4	70,6% (-2,7%)	92,8% (7,1%)
Chan et al. *	CA19-9, Ca125 and LAMC2	Logistic regression: CA19.9 + 1.13 · CA125 + 0.143 · LAMC2	83,2% (10,4%)	81,54% (0%)
Gu et al. *	CA19-9, CEA, Ca125 and CA242	"Joint detection test"	90,4% (7,7%)	93,8% (35,2%)
Hogendorf et al. 2017	CA19-9, Ca125 and Ca15-3	Logistic regression: 0,253 + 1,039 ∗ CA19-9 + 1,003 ∗ CA125 + 1,048 ∗ CA15-3.	81,2% (22,23%)	63,1% (-16,25%)
Hogendorf et al. 2018	Ca125 and GDF-15	Logistic regression: 0,23+1,001*GDF-15+1,07*Ca125.	80% (22,86%)	80,95% (-9,55%)

Table 4. The summary of diagnostic performance of each proposed combined test. The difference between test's sensitivity, specificity and that of CA19-9 is given in the brackets.

Bold font marks the studies proposing a combined test that outperforms CA19-9 in terms of the overall performance.

"" marks the studies with the combined test outperforming CA19-9 in terms of both sensitivity and specificity.*

σ Study by Wang et al. propose a combined test of CA19-9 and CA72-4

Supplementary Tables

Ca125

Variable	Sensitivity Intercept	p	Sensitivity (Variable)	p	Specificity Intercept	p	Specificity (Variable)	p
ut-off point estimation	1,5 (1,04-1,94)	< 0,0001	0,915 (0,137-1,69)	**0,021**	-3,23 (-3,88--2,58)	< 0,0001	0,78 (-0,29-1,84)	0,15
udy location	1,66 (0,94-2,37)	< 0,0001	0,33 (-0,6 – 1,27)	0,7	-3,48 (-4,32 - -2,65)	< 0,0001	0,87 (-0,2 - -1,94)	0,11
Study type	1,6 (0,77-2,44)	< 0,0001	0,28 (-0,77-1,29)	0,58	-3 (-3,98--2,02)	< 0,0001	0,01 (-1,2-1,21)	0,99
Biomarker essment type*	1,55 (0,88-2,22)	< 0,0001	0,48 (-0,42-1,38)	0,3	-3,29 (-4,13 - -2,45)	<0,0001	0,56 (-0,53-1,65)	0,31
AC Prevalence (>60%)	1,34 (0,76-1,92)	< 0,0001	0,8 (0,03-1,57)	**0,04**	-3,3 (-4,1--2,49)	< 0,0001	0,55 (-0,52-1,62)	0,32

Supplementary Table 1. The coefficients for the bivariate models used for Ca125 meta-regression

CA19-9

Variable	Sensitivity Intercept	p	Sensitivity (Variable)	p	Specificity Intercept	p	Specificity (Variable)	p
ut-off point estimation	3,21 (2,76-3,66)	< 0,0001	-1,17 (-1,8 - -0,54)	**<0,0001**	-2,79 (-3,34 - - 2,4)	< 0,0001	-0,68 (-1,67-0,312)	0,18
udy location	3,28 (2,6-3,95)	< 0,0001	-0,91 (-1,74 - -0,08)	**0,032**	-2,53 (-3,1 - -1,97)	< 0,0001	-0,85 (-1,65 - -0,05)	**0,038**
Study type	3,16 (2,46-3,85)	< 0,0001	-0,72 (-1,61 - 0,17)	0,11	-2,6 (-3,17 - -2,03)	< 0,0001	-0,83 (-1,71 - -0,05)	0,064
Biomarker ssment type*	3,32 (2,69-3,96)	< 0,0001	-0,97 (-1,76 - -0,18)	**0,016**	-2,96 (-3,71 - -2,21)	<0,0001	-0,12 (-1.11 – 0,87)	0,82
PDAC Prevalence (>60%)	3,2 (2,47-3,93)	< 0,0001	-0,76 (-1,67 – 0,15)	0,1	-2,61 (-3,2 – -2,01)	< 0,0001	-0,73 (-1,59 – 0,13)	0,097

Supplementary Table 2. The coefficients for the bivariate models used for CA19-9 meta-regression

Supplementary Figure Legends

Supplementary Figure 1. Funnel plot for publication bias analysis of CA19-9

Supplementary Figure 2. Funnel plot for publication bias analysis of Ca125

Figures

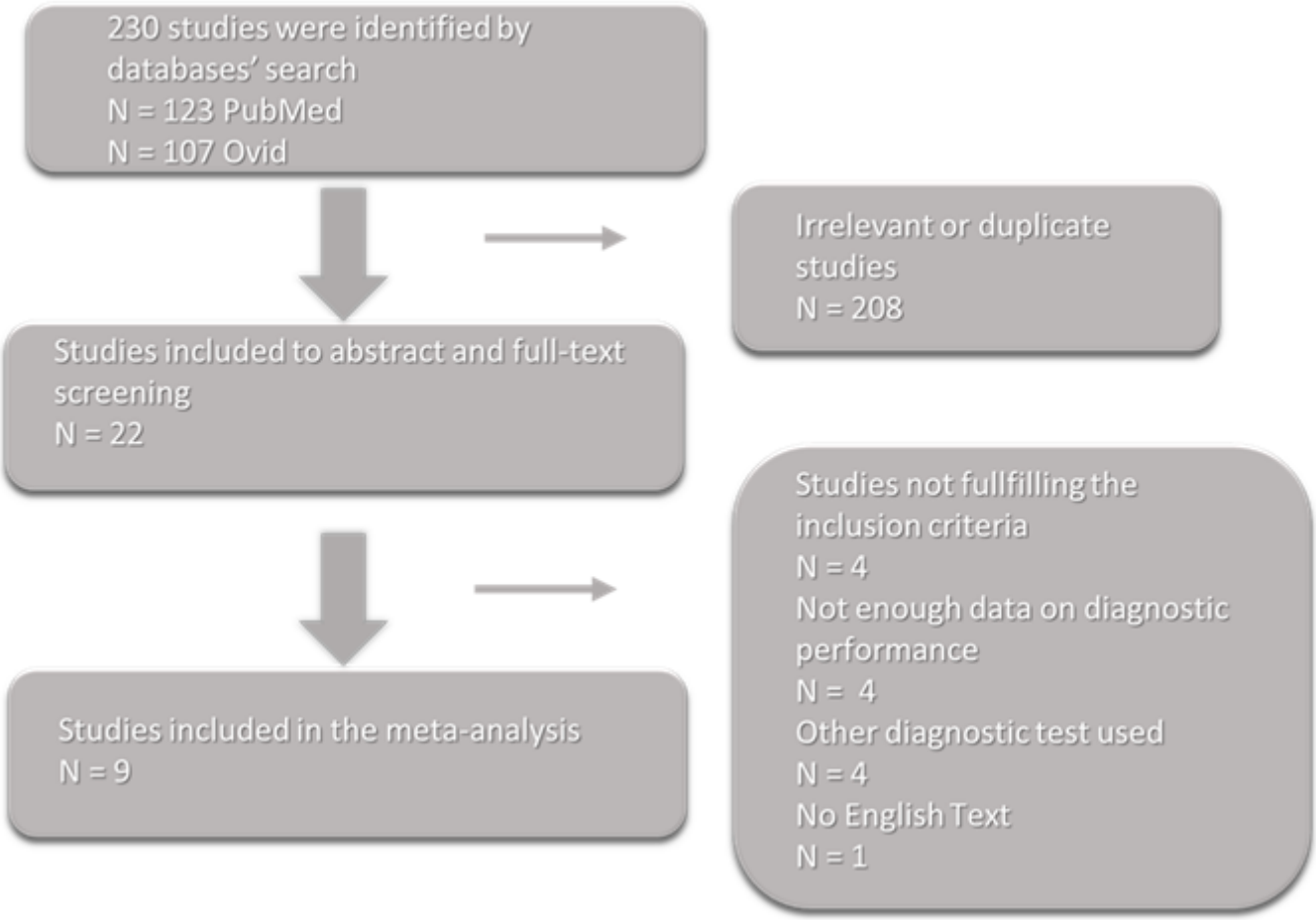

Figure 1

Flow diagram of meta-analysis

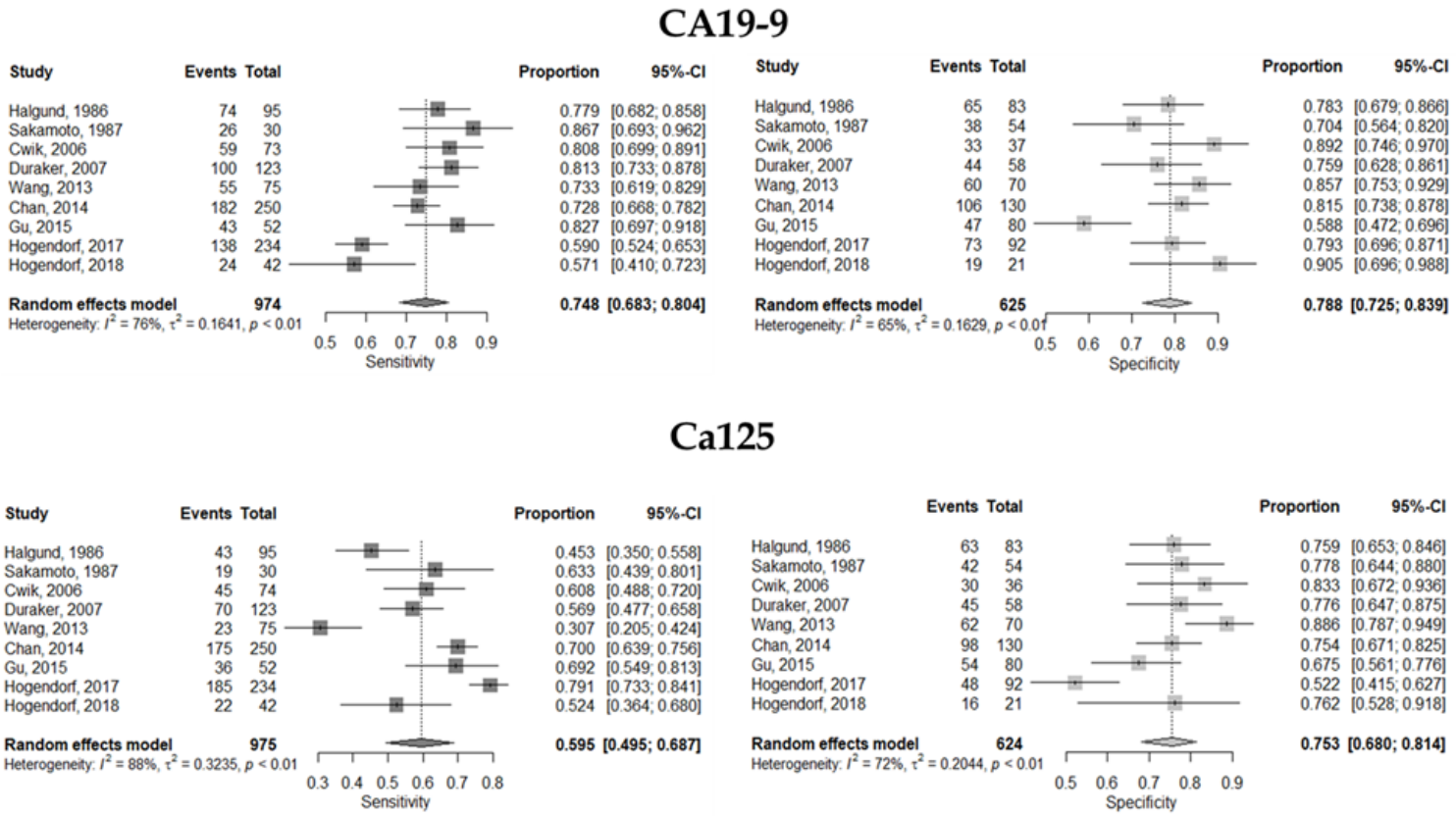

Figure 2

Forest plots depicting biomarkers' sensitivity and sensitivity. Left panel: sensitivity, right panel: specificity.

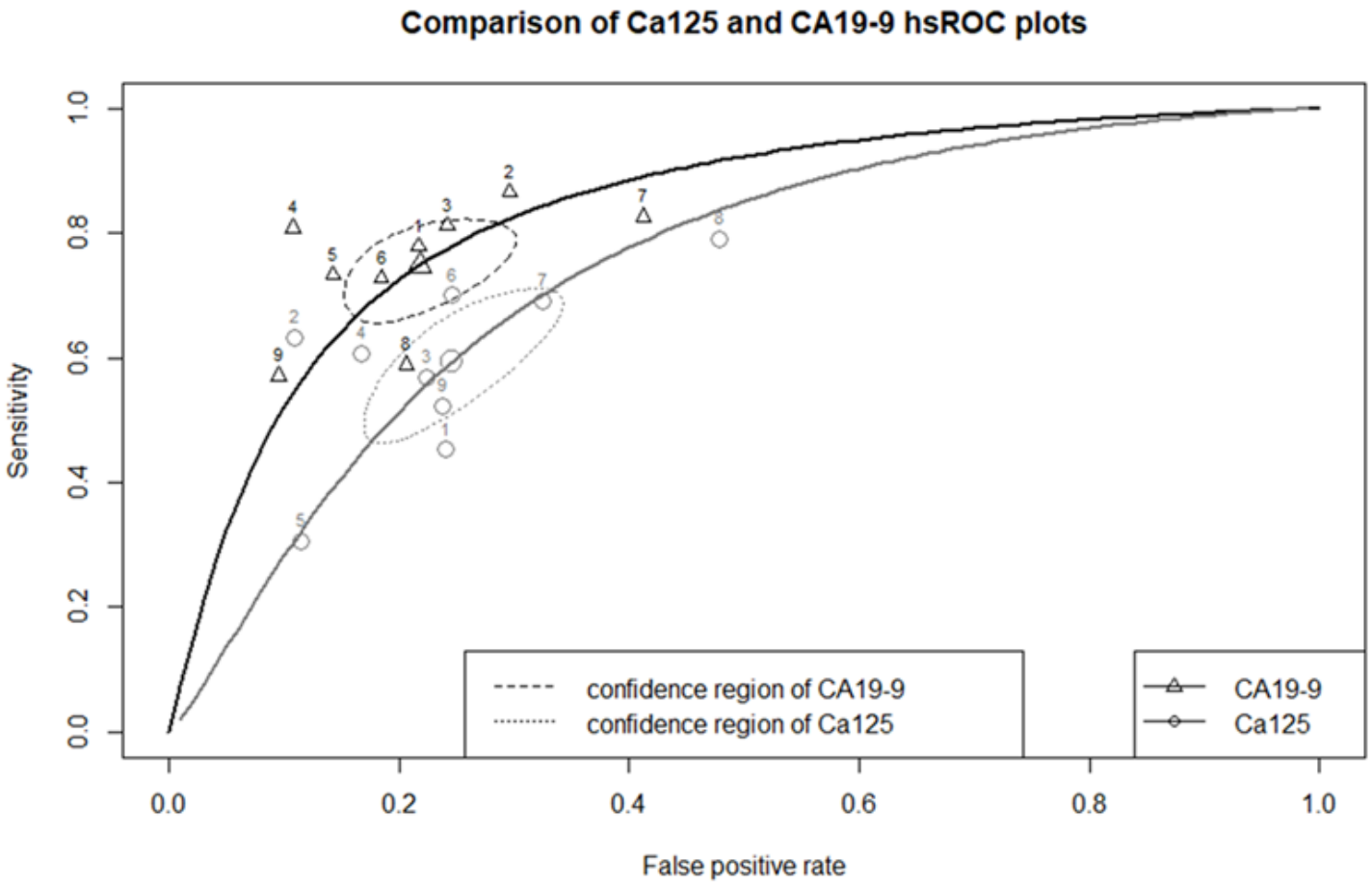

Figure 3

Comparison of Ca125 and CA19-9 HSROC curves. The numerical labels of points has been assigned to each study in a chronological manner (e.g. 1- Halgund, 2- Sakamoto etc.)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- PRISMA2009CA199Ca125Metaanalysis.pdf
- SuppFigure2funnelplot.jpg
- SuppFigure1funnelplot.jpg