PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Soluble urokinase plasminogen activator receptor (suPAR) as a prognostic marker of mortality in healthy, general, and patient populations: protocol for a systematic review and meta-analysis
AUTHORS	Petersen, Jens Emil; Kallemose, Thomas; Barton, Karen; Caspi, A; Rasmussen, Line Jee Hartmann

VERSION 1 – REVIEW

REVIEWER	Bjorn Meijers KU Leuven and University Hospitals Leuven Herestraat 49 B-3000 Leuven Belgium
REVIEW RETURNED	23-Dec-2019

GENERAL COMMENTS	This reviewer agrees with the authors that a meta-analysis of suPAR and outcomes is much needed. In the search strategy it is recommended to add "soluble urokinase receptor" It is now generally accepted that kidney function is an important determinant of blood concentrations of suPAR. While it is mentioned to perform subgroup analyses for presence of chronic kidney disease (page 25), apparently markers of kidney function are not collected in a structured way. It is advisable to collect (i) whether kidney function is determined, (ii) whether kidney function is determined at the day of sampling for suPAR, (iii) what marker of kidney function is used and how it is measured, (iv) whether outcomes have been adjusted for kidney function, and if so, how adjustment has been done (v) which covariate (creatinine, (measured or estimated) creatinine clearance, eGFR, mGFR) has been entered into multivariate models. In the protocol it is assumed that suPAR has a normal distribution of log-normalized distribution. Additional steps will be taken if this would be not the case. It is advisable to include subgroup analyses for healthy individuals vs. populations of diseased individuals

REVIEWER	Nils A. Sörensen University Heart and Vascular Center Hamburg Department of Cardiology Germany
REVIEW RETURNED	23-Dec-2019

GENERAL COMMENTS	Petersen et al. plan to perform a Metaanalysis on the prognostic value of the biomarker "Soluble urokinase plasminogen activator receptor (suPAR)". A comparable analysis has not been published before and I do believe it is worth the effort. However, some adjustments should be made: - The planned search strategy and, most important, the tools to detect the risk of bias in the included studies are well discribed and seem appropriate. However, it is important, that disagreement on the risk of bias by the two reviewers, which needs resolving by a third reviewer will be displayed in the final manuscript, this should be incorporated in the methods section. - An important aim, of this metaanalysis should be the development of concrete recommendations for clinicians on how suPAR could be used in the different discribed clinical settings (healthy, general, and various patient groups). So far, the protocol focuses on the simple association of suPAR-levels and mortality. The authors should try to establish suPAR cutoff levels in the respective setting so one can identify patients at very low risk to high risk based on their suPAR level. A strategy on how this could be achieved should be added in the methods section. - Given, the high value of suPAR for cardiovascular diseases, I recommend to not only adjust and compare suPAR to other inflammatory biomarkers, but include cardiovascular biomarkers (e.g. NT-proBNP, troponin) in the analysis.

REVIEWER	Ivan Cavero-Redondo Universidad de Castilla-La Mancha
REVIEW RETURNED	14-Jan-2020

GENERAL COMMENTS	I must congratulate the authors for this manuscript entitled: "suPAR as a prognostic marker of mortality in healthy, general, and patient populations: protocol for a systematic review and meta-analysis", for two fundamental reasons: First, the research question is novel and interesting, with the results of this review being important for future research and for the inclusion of this biomarker in clinical practice. Second, the authors offer a master lesson on how to conduct a systematic review, making the methodology reproducible for any reader of this article. I only have minor comments that I would like the authors to respond in order to improve some aspects of their work. 1. The authors intend to include the results obtained from the control groups of randomized clinical trials. While it is true, that the control group does not receive any treatment and we could consider it an observational group, it could be that this control group was blinded and did not know if they received the treatment or not, and a placebo effect could occur. This is an important limitation because although this mortality may not be so important, in the suPAR biomarker, this placebo effect could alter its levels. 2. The authors intend to make an analysis in the human patient population, but this group is very heterogeneous. I think it makes no sense to include patients with cardiometabolic pathology or infectious pathologies in the same group. The authors should further limit the human patient population group to a homogeneous group of patients. 3. Although the RR, OR and HR have the same interpretation, it does not seem logical that they can be combined without having turned them all into the same measure of association. In the case

of the RR and the OR, it is very likely that they have been calculated using a 2x2 table (although they can be obtained through regression), while in the case of the HR it is necessary to perform a Cox regression and the time variable is important. I recommend that the authors explain how they intend to combine these three measures of association into one.

4. Although I2 and Chi2 are usually used as measures of heterogeneity, they are actually measures of inconsistency. In my opinion, I2 is more understandable for the reader, since it provides a percentage of variability between studies and is calculated from Chi2. I encourage authors to maintain I2 and replace Chi2 with tau2 as a measure of heterogeneity.

5. Setting the I2 value> 50% is inaccurate. Authors should include the levels of heterogeneity established by the last version of the Cochrane Handbook to establish the presence and level of inconsistency/heterogeneity

6. Regarding the use of ROC curves, I believe that the authors should use a different approach, performing an analysis using HSROC and an analysis of the diagnostic Odds Ratio (dOR), this would substantially improve the article

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Bjorn Meijers

Institution and Country:

KU Leuven and University Hospitals Leuven Herestraat 49

B-3000 Leuven

Belgium

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

This reviewer agrees with the authors that a meta-analysis of suPAR and outcomes is much needed.

In the search strategy it is recommended to add "soluble urokinase receptor"

Thank you.

We have added the term "soluble urokinase receptor" under the section "Search Strategy", thus the text was changed to:

> "The following terms will be used to search the electronic databases in addition to other related terms for the concepts of "suPAR" and "mortality": "suPAR" or "soluble urokinase plasminogen

activator receptor" or "soluble urokinase-type" or "soluble urokinase receptor" or "uPAR" AND "mortality" or "death" or "fatality""

The term was already included in the full search strategy (Appendix 1), which we have now emphasized by rephrasing the last sentence of the Search strategy-paragraph to:

"The full PubMed search and search terms are shown in **Appendix 1**."

It is now generally accepted that kidney function is an important determinant of blood concentrations of suPAR. While it is mentioned to perform subgroup analyses for presence of chronic kidney disease (page 25), apparently markers of kidney function are not collected in a structured way.

It is advisable to collect (i) whether kidney function is determined, (ii) whether kidney function is determined at the day of sampling for suPAR, (iii) what marker of kidney function is used and how it is measured, (iv) whether outcomes have been adjusted for kidney function, and if so, how adjustment has been done (v) which covariate (creatinine, (measured or estimated) creatinine clearance, eGFR, mGFR) has been entered into multivariate models.

We thank Dr. Meijers for this suggestion.

We acknowledge the potential effect of kidney function on suPAR concentration and have therefore included the recommended markers of kidney function (creatinine, creatinine clearance, and glomerular filtration rate) in the data collection. These measurements will be included in our subgroup analyses. Furthermore, our meta-analysis plan already includes a maximally adjusted group which includes associations adjusted for kidney function if included in the multivariate analyses.

To reflect this, the paragraphs in the following sections have been modified as shown:

Eligibility Criteria:

"Comparators: We will investigate the unadjusted and adjusted prognostic value of suPAR, i.e., without and with adjustments for other prognostic factors, e.g., conventional risk factors (such as age, sex, smoking, and chronic disease), inflammatory biomarkers (such as CRP, white blood cells, and IL-6), or kidney function (such as creatinine and glomerular filtration rate)."

Data items:

"[…]; (6) control characteristics (conventional risk factors, e.g., age, sex, smoking, and chronic diseases; other inflammatory biomarkers, e.g., C-reactive protein (CRP), white blood cells,

cytokines, fibrinogen; and kidney function, e.g., creatinine (measured or estimated), creatinine clearance, glomerular filtration rate (measured or estimated)); [...]"

Subgroup analyses and meta-regression:

"adjustment for kidney function;" has been added to the list of subgroup analyses.

In the protocol it is assumed that suPAR has a normal distribution of log-normalized distribution. Additional steps will be taken if this would be not the case. It is advisable to include subgroup analyses for healthy individuals vs. populations of diseased individuals

We agree that this is an important subgroup analysis, and the approach is described in the section "Subgroup analyses and meta-regression", where subgroup analyses will be based on population type (healthy/general population vs. patient types, e.g., cardiovascular disease, cancer, chronic kidney disease, infectious disease, critical illness, acute care).

Reviewer: 2

Reviewer Name: Nils A. Sörensen

Institution and Country:

University Heart and Vascular Center Hamburg Department of Cardiology Germany Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below

Petersen et al. plan to perform a Metaanalysis on the prognostic value of the biomarker "Soluble urokinase plasminogen activator receptor (suPAR)". A comparable analysis has not been published before and I do believe it is worth the effort. However, some adjustments should be made:

- The planned search strategy and, most important, the tools to detect the risk of bias in the included studies are well discribed and seem appropriate. However, it is important, that disagreement on the risk of bias by the two reviewers, which needs resolving by a third reviewer will be displayed in the final manuscript, this should be incorporated in the methods section.

Thank you.

We agree with this suggestion and have updated the Methods section to include a strategy to document and report disagreements in the quality assessment:

> "Any disagreements will be resolved by consensus, or if necessary by a third author, and a log of these will be included as an appendix in the final manuscript. No study will be excluded based on the results of risk of bias assessment. We will compute graphic representations of potential bias for the final manuscript. In the meta-analysis, subgroup analyses will be performed based on the risk of bias (QUIPS; low, moderate, or high risk of bias). The adapted QUIPS tool will be provided as an appendix in the final manuscript along with the log of disagreements."

- An important aim, of this metaanalysis should be the development of concrete recommendations for clinicians on how suPAR could be used in the different discribed clinical settings (healthy, general, and various patient groups). So far, the protocol focuses on the simple association of suPAR-levels and mortality. The authors should try to establish suPAR cutoff levels in the respective setting so one can identify patients at very low risk to high risk based on their suPAR level. A strategy on how this could be achieved should be added in the methods section.

This is an excellent suggestion. However, since this systematic review and meta-analysis will be based on risk estimates of study populations, the lack of individual-level information on suPAR and mortality hinders the definition of specific suPAR concentration cut-offs and is outside the scope of this study.

But we have updated our analysis section to include a strategy that allows for establishing the sensitivity (detection rate) at a 10% false positive rate.

> "As described for CRP by Hemingway et al.,[69] we will attempt to calculate the detection rate (sensitivity) at different false positive rates from 0 to 100 by constructing the log-normal distributions of suPAR separately for those who survived and those who died. From this we will obtain a ROC curve and report the c-statistic. Pooled estimates of both the c-statistic and detection rate of suPAR's discriminative ability for predicting mortality will be obtained by random effects meta-analysis of the study-specific c-statistics and detection rates. Confidence intervals and a 10% false positive rate will be reported."

Moreover, as mentioned in the response to Reviewer 3, comment #6, we will perform HSROC analysis. Here, we will attempt to evaluate cut-offs based on Q and diagnostic odds ratio (measures of overall accuracy based on sensitivity and specificity), given that the thresholds used in each study are similar enough for this purpose.

- Given, the high value of suPAR for cardiovascular diseases, I recommend to not only adjust and compare suPAR to other inflammatory biomarkers, but include cardiovascular biomarkers (e.g. NT-proBNP, troponin) in the analysis.

The meta-analysis plan already includes a maximally adjusted group which includes associations adjusted for cardiovascular biomarkers if included in the multivariate analyses of the individual studies.

Reviewer: 3

Reviewer Name: Ivan Cavero-Redondo

Institution and Country: Universidad de Castilla-La Mancha

Please state any competing interests or state 'None declared': None declare

Please leave your comments for the authors below

I must congratulate the authors for this manuscript entitled: "suPAR as a prognostic marker of mortality in healthy, general, and patient populations: protocol for a systematic review and meta-analysis", for two fundamental reasons: First, the research question is novel and interesting, with the results of this review being important for future research and for the inclusion of this biomarker in clinical practice. Second, the authors offer a master lesson on how to conduct a systematic review, making the methodology reproducible for any reader of this article. I only have minor comments that I would like the authors to respond in order to improve some aspects of their work.

1. The authors intend to include the results obtained from the control groups of randomized clinical trials. While it is true, that the control group does not receive any treatment and we could consider it an observational group, it could be that this control group was blinded and did not know if they received the treatment or not, and a placebo effect could occur. This is an important limitation because although this mortality may not be so important, in the suPAR biomarker, this placebo effect could alter its levels.

Thank you.

As Dr. Cavero-Redondo points about, the placebo effect could influence the suPAR levels and mortality in randomized control studies. To control for this, we have updated the protocol to include a subgroup analysis comparing control groups from randomized controlled trials against the rest of the studies.

> "Subgroup analyses will be used to explore possible sources of heterogeneity, and univariate random effects meta-regression will be performed based on the following: study design (cohort, case-control, randomized controlled trials); […]"

However, suPAR and mortality would likely be similarly affected in the presence of the placebo effect, and we do not expect the association between the two to change

2. The authors intend to make an analysis in the human patient population, but this group is very heterogeneous. I think it makes no sense to include patients with cardiometabolic pathology or infectious pathologies in the same group. The authors should further limit the human patient population group to a homogeneous group of patients.

The inclusion of the very diverse group of patients is part of our ambitious aim to evaluate the association between suPAR and mortality in as wide a population sample as possible. To deal with the heterogeneity in the included populations, we will stratify the populations into disease specific subgroups and perform separate subgroup analyses as described in the section "Subgroup analyses and meta-regression" (as shown below), and compare these to the primary analysis.

> "Subgroup analyses will be used to explore possible sources of heterogeneity, and univariate random effects meta-regression will be performed based on the following: [...] population type (healthy/general population vs. patient types, e.g., cardiovascular disease, cancer, chronic kidney disease, infectious disease, critical illness, acute care); [...]"

3. Although the RR, OR and HR have the same interpretation, it does not seem logical that they can be combined without having turned them all into the same measure of association. In the case of the RR and the OR, it is very likely that they have been calculated using a 2x2 table (although they can be obtained through regression), while in the case of the HR it is necessary to perform a Cox regression and the time variable is important. I recommend that the authors explain how they intend to combine these three measures of association into one.

We agree with Dr. Cavero-Redondo that combining different measures of risk estimates gives rise to certain limitations. This approach will only give us the direction of an association and does not allow us to provide an absolute, pooled effect size, since the different risk estimates are reported on different scales.

Relative risks with 95-99% CIs will be used as the common measure of association across studies. Relative risks, odds ratios, and hazard ratios will be assumed to approximate the same measure of relative risk. For the primary analysis all study outcome measures (e.g., RR, OR, and HR) will be pooled as a single measure, and all available studies will be included, regardless of population. If a study has multiple versions of the same model with different adjustments, the model with most adjustments will be included.

To clarify, we have added the following to the section "Data synthesis":

> "Relative risks with 95-99% CIs will be used as the common measure of association across studies. RRs, ORs, and HRs will be assumed to approximate the same measure of relative risk."

To account for the limitations associated with combining these measures, we will in addition to the main analysis I) conduct separate meta-analyses on the most long-term, most adjusted outcome stratified on individual outcome measure, e.g., RR, OR, and HR (as described on p. 24 in the protocol), and II) we will conduct individual meta-regressions to account for the heterogeneity across methods of reporting outcomes. Thus, we have added the following to the list of subgroup analyses and meta-regressions:

> "[…]; reported relative risk estimates (e.g., RR, OR, HR); […]"

4. Although I2 and Chi2 are usually used as measures of heterogeneity, they are actually measures of inconsistency. In my opinion, I2 is more understandable for the reader, since it provides a percentage of variability between studies and is calculated from Chi2. I encourage authors to maintain I2 and replace Chi2 with tau2 as a measure of heterogeneity.

We thank Dr. Cavero-Redondo for making this clarification.

We have revised the protocol according to this suggestion, replacing Chi2 with Tau2. Thus, the text was changed to:

> "Statistical heterogeneity among studies will be evaluated using the Tau^2 and I^2 statistic (where I^2 of 30-60% will be interpreted to indicate moderate heterogeneity and I^2 >50% to indicate substantial heterogeneity across studies[71])."

> *71. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ WV (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.*

5. Setting the I2 value> 50% is inaccurate. Authors should include the levels of heterogeneity established by the last version of the Cochrane Handbook to establish the presence and level of inconsistency/heterogeneity 6.

Thank you for this comment. We have revised the description accordingly, using directions provided by the Cochrane Handbook (*Higgins JPT, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook*). As shown in the previous comment, we have changed the text to:

"Statistical heterogeneity among studies will be evaluated using the Tau^2 and I^2 statistic (where I^2 of 30-60% will be interpreted to indicate moderate heterogeneity and I^2 >50% to indicate substantial heterogeneity across studies[71])."

We intend to evaluate the heterogeneity and interpret this based on the Cochrane Handbook; however, despite a substantial level of heterogeneity (which is expected due to the mix of healthy and patient populations as previously described), we still intend to run the pooled models although we further plan on exploring the heterogeneity and evaluating this for subgroups as detailed in the protocol.

6. Regarding the use of ROC curves, I believe that the authors should use a different approach, performing an analysis using HSROC and an analysis of the diagnostic Odds Ratio (dOR), this would substantially improve the article

Thank you for suggesting this approach. We have changed the approach accordingly and added the following text:

"Analysis of the predictive value of suPAR for mortality will be done by hierarchal summary receiver operation characteristic (HSROC) model curves. From this, SROC curves with AUCs, Qs, and diagnostic odds ratios (DORs) will be produced."

To calculate these, we need true positive (TP), false positive (FP), true negative (TN), false negative (FN) frequencies. These have been added to the list of data items to be collected:

"[…] and true positive (TP), false positive (FP), true negative (TN), and false negative (FN) frequencies as well as AUCs for ROC curves); […]"

VERSION 2 – REVIEW

REVIEWER	Bjorn Meijers KU Leuven and UZ Leuven Belgium
REVIEW RETURNED	12-Feb-2020

GENERAL COMMENTS	The authors have addressed my concerns

REVIEWER	Nils A. Sörensen University Heart and Vascular Center, Hamburg, Germany
REVIEW RETURNED	27-Feb-2020

GENERAL COMMENTS	The authors have addressed my comments appropriately.

REVIEWER	Iván Cavero Redondo Universidad de Castilla-La Mancha
REVIEW RETURNED	12-Feb-2020

GENERAL COMMENTS	The authors have improved the manuscript. Congrats