# Accuracy of recommended sampling and assay methods for the determination of plasma free and urinary fractionated metanephrines in the diagnosis of pheochromocytoma and paraganglioma: A systematic review

**Roland Därr**[1], **Matthias Kuhn**[2], **Christoph Bode**[1], **Stefan R. Bornstein**[3], **Karel Pacak**[4], **Jacques W.M. Lenders**[3,5], and **Graeme Eisenhofer**[3,6]

[1]Department of Cardiology and Angiology I, University Heart Centre Freiburg, Freiburg, Germany
[2]Institute for Medical Informatics and Biometry, Faculty of Medicine Carl Gustav Carus at the TU Dresden, Dresden, Germany [3]Department of Medicine III, University Hospital Carl Gustav Carus at the TU Dresden, Dresden, Germany [4]Program in Reproductive and Adult Endocrinology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA [5]Department of Internal Medicine, Radboud University Medical Center, Nijmegen, The Netherlands [6]Institute of Clinical Chemistry and Laboratory Medicine, University Hospital Carl Gustav Carus at the TU Dresden, Dresden, Germany

## Abstract

**Purpose:** To determine the accuracy of biochemical tests for the diagnosis of pheochromocytoma and paraganglioma (P/PGL).

**Methods:** A search of the PubMed-database was conducted for English-language articles published between October 1958 and December 2016 on the biochemical diagnosis of P/PGL using immunoassay (IA) methods or high-performance liquid chromatography (HPLC) with coulometric/electrochemical (AC/EC) or tandem mass-spectrometric (MS/MS) detection for measurement of fractionated metanephrines in 24-h urine collections (uMNs) or plasma free metanephrines (pMNs) obtained under seated or supine blood sampling conditions.

**Results:** Application of the Standards for Reporting of Diagnostic Studies Accuracy (STARD) Group criteria yielded 23 suitable articles. Summary receiver operating characteristic (sROC) analysis revealed sensitivities/specificities (SE/SP) of 94%/93% and 91%/93% for measurement of pMNs and uMNs using HPLC or IA methods, respectively. Partial areas under the curve (AUCs) were 0.947 versus 0.911. Irrespective of the analytical method, SE was significantly higher for supine compared to seated sampling, 95% versus 89% (p<0.02), while SP was significantly higher for supine sampling compared to 24-h urine, 95% versus 90% (p<0.03). Partial AUCs were 0.942, 0.913, and 0.932 for supine sampling, seated sampling, and urine. Test accuracy increased linearly

Corresponding author: Roland Därr, MD, Department of Cardiology and Angiology I, University Heart Centre Freiburg, Hugstetter Straße 55, 79106 Freiburg, Germany, Phone: +49-(0)761 270-34411, Fax: +49-(0)761 270-34412, Roland.wolfgang.daerr@uniklinik-freiburg.de.

Compliance with Ethical Standards

Author Manuscript

from 90% to 93% for 24-h urine at prevalence rates of 0.0–1.0, decreased linearly from 94% to 89% for seated sampling and was constant at 95% for supine conditions.

**Conclusions:** Current tests for the biochemical diagnosis of P/PGL show excellent diagnostic accuracy. Supine sampling conditions and measurement of pMNs using HPLC with AC/EC or MS/MS detection provides the highest accuracy at all prevalence rates.

## Keywords

Pheochromocytoma; paraganglioma; metanephrines; immunoassay; high performance liquid chromatography; diagnostic accuracy

## 1. Introduction

Pheochromocytoma (P) and paragangliomas (PGL) are rare neuroendocrine tumors arising from chromaffin cells respectively of the adrenal medulla or associated with sympathetic ganglia. These tumors are characterized by an excessive production of catecholamines. They are one of a few curable causes of hypertension, but can also mimick numerous other diseases and remain unrecognized over prolonged periods of time [1]. Also, sudden bursts of catecholamine release may have devastating clinical consequences [2]. Thus, precise and sensitive diagnostic tests are of paramount importance for the reliable diagnosis of P/PGL.

It is now commonly agreed that measurements of plasma free metanephrines (pMNs) [3,4] or urinary fractionated metanephrines (uMNs) [5,6] provide the best available biochemical tests for diagnosis of P/PGL, where pMNs are defined as the unconjugated 3-O-methylated primary degradation products metanephrine (MN) and normetanephrine (NMN) of the catecholamines adrenaline and noradrenaline, and uMNs reflecting the sum of the respective 3-O-methylated free and sulfate-conjugated metabolites [7]. Simultaneous measurement of the 3-O-methylated metabolite of dopamine, methoxytyramine (MTY), has been advocated in particular for the diagnosis of P/PGL due to an underlying mutation of the succinate dehydrogenase B (SDHB) gene which is associated with an increased rate of metastatic disease [8]. However, the added value of MTY for identifying P/PGL in an unselected screening population is less clear [9,10], and in fact, the vast majority of studies on the diagnostic efficiency of pMNs or uMNs have not included MTY, apart from a few exceptions [9,11,10].

Likewise, much of the discussion in the past has revolved around the optimal assay method for quantification of pMNs and uMNs. Meanwhile considerable evidence has accumulated in support of the superiority of high performance liquid chromatography (HPLC) coupled with coulometric/electrochemical (AC/EC) or tandem mass spectrometric (MS/MS) detection over immunoassay (IA) techniques (RIA, EIA, ELISA) that appear to underestimate particularly pMNs [12,13]. On the other hand the precision of HPLC methods in particular if coupled with MS/MS detection obviously has its price. Capital costs of instrumentation, a high level of operator expertise and the need for in-house validation are in stark contrast with the technical, personnel and financial requirements of IA methods [14]. Against this background the currently recommended best practice procedure for the biochemical diagnosis of P/PGL, i.e. blood sampling in resting-supine position after an

overnight fast when MTY is included [15] and measurement of pMNs using HPLC with AC/EC or MS/MS detection may not always be followed as suggested [6,16]. In this case, however, interpretation of test results need to take into account potential losses in diagnostic certainty associated with the respective assay method and sampling procedure being used and an increased necessity for confirmatory testing.

The present systematic review was performed to examine in detail the diagnostic accuracy of measurements of pMNs and uMNs by HPLC or IA techniques and the impact of blood sampling conditions. As a result a simple decision guide is provided for selecting the most suitable test procedure for the biochemical diagnosis of P/PGL depending on the anticipated risk of disease.

## 2. Methods

### 2.1. Data sources, study selection and conceptual design

For the present investigation the PubMed-database was searched for all English language articles on the diagnostic accuracy of measurement of pMNs and uMNs in the biochemical diagnosis of P/PGL published between October 1958 and December 2016 with an emphasis on articles predominantly of the last three decades reporting the use of highly sensitive IA systems or HPLC coupled with AC/EC or MS/MS detection. Search terms were "Metanephrine"[MAJR] and "Humans"[MeSH Terms]. Additionally, the reference lists of pertinent articles including a single 2004 systematic review on the diagnostic performance of measurement of pMNs [17] were evaluated by a hand search. Titles and abstracts of the articles identified were scrutinized by two independent raters. Publications not in English language, letters, editorials, comments, review articles, and single case reports were not considered. Of the full-text publications studies on differing topics, studies with biased trial population or systematic limitations, and methodology studies were excluded. In 1 case of consecutive publishing from the same research group with apparently overlapping study populations and broadly convergent results only the most recent study describing the largest number of test subjects was included. The conceptual design of the investigation followed the recommendations of the PRISMA (Preferred Reporting Items for Systematic Reviews and Metaanalyses) statements [18].

### 2.2. Quality of methodology and reporting of included studies

To assess the quality of methodology and reporting of the studies deemed appropriate for inclusion in the present systematic review by both raters the criteria developed by the Standards for Reporting of Diagnostic Studies Accuracy (STARD) Group were applied [19,20], and the degree of consensus reached between the two raters on the 25-items STARD checklist was determined by assigning one vote per applicable item (n=21) to each rater.

### 2.3. Statistical Analysis

"R" version 3.1.1 (The R Foundation for Statistical Computing, Vienna, Austria) and the "mada" package were used to construct summary receiver operating characteristic (sROC) curves from sensitivity and specificity values derived from individual studies as described by Reitsma *et al* [21]. To avoid zero frequencies in calculations a continuity correction of 0.5

was applied to all observed frequencies. Specificities and sensitivities were compared by fitting a bivariate meta-regression on sensitivity and false positive rate with fixed effects for each of the factors analytical method (i.e. IA or HPLC) and sample/sampling condition (i.e. blood obtained in a supine or seated position or 24-h urine collections). For transformation to the logit link respective parameters were set at 0.5 and 1.5 for continuity corrected sensitivities and false positive rates based on Akaike information criterion. Significance of factors was assessed using the likelihood ratio test and subsequent posthoc pairwise comparison with Wald test, if applicable. Accuracy was defined as the sum of the true positive and true negative test results divided by the total population of a selected study. Test accuracy as a function of pretest likelihood (i.e. variable disease prevalence levels between 0.0 and 1.0) was calculated for supine and seated plasma sampling and 24-h urine collection according to the following equation: $ACC = SE \times PRE + SP \times (1-PRE)$, where ACC denotes accuracy, SE sensitivity, SP specificity, and PRE disease prevalence. Cohen's Kappa ($\kappa$-value) was determined as a measure of interrater agreement.

## 3.    Results

### 3.1.    Study Selection and quality assessment according to the STARD criteria

Search of the PubMed database and manual evaluation of the reference lists of relevant articles yielded a total of 284 results. Exclusion of studies on methodology (n=79), studies dealing with different topics (n=83), studies with biased trial population (n=19), studies not in English language (n=15), studies with systematic limitations (n=13), letters to the editor (n=24), case reports (n=11), review articles (n=10), editorials (n=6), and comments (n=1) resulted in 23 full-text studies for further consideration (Fig. 1). Assessment of the quality of methodology and reporting of these studies showed a median score for compliance with 21 applicable criteria of the STARD checklist of 33.0 (range 22.0–37.0). The median $\kappa$-value for interrater consent was 0.828 (range 0.525–0.920).

### 3.2.    Characteristics of included studies

The diagnostic performance of measurements of pMNs and uMNs in the detection of P/PGL, the respective assay methods, plasma sampling conditions, total numbers of subjects, and numbers of patients and controls from the 23 included studies in chronological order according to year of publication are summarized in Table 1. The number of trial subjects of all studies totaled 8,173, including 1,039 patients with P/PGL and 7,134 controls in whom P/PGL had been ruled out. The median number of subjects per study was 184 (range 31–1,819). IA techniques for measurements of pMNs and/or uMNs were used in 11 studies (1,530 controls; 350 patients) and HPLC with AC/EC or MS/MS detection in 15 (6,019 controls; 794 patients). Blood samples were taken with the subjects in a supine position in 10 studies (3,360 controls; 588 patients) and in a seated position also in 10 (1,364 controls; 285 patients). Measurements of uMNs was carried out in 11 studies (3,698 controls; 467 patients). Supine blood sampling conditions, blood sampling in a seated position, and measurements of uMNs in 24-h urine collections were reported in 38.5%, 38.5%, and 23.1% of studies using IA analysis methods and in 28.6%, 33.3%, and 38.1% of studies using HPLC. Percentage differences were not statistically significant. Sensitivities and specificities among studies ranged from 76% to 100% and from 69% to 100%, respectively.

### 3.3.    Data Synthesis

**3.3.1.    SROC curves for IA and HPLC analysis techniques—**The diagnostic performance of measurements of pMNs and uMNs in 24-h urine collections from studies using either IA techniques or HPLC was evaluated by sROC analysis (Fig. 2). The respective point estimators of sensitivities and specificities from the resultant sROC curves were 91% (CI 87–94%) and 93% (CI 89–96%) for IA compared with 94% (CI 91–96%) and 93% (CI 90–95%) for HPLC. These differences were statistically not significant. Values of the area under the curve (AUC) restricted to observed false-positive rates (FPRs) and normalized were 0.911 for IA and 0.947 for HPLC.

**3.3.2.    SROC curves for blood sampling in the supine and seated position and for 24-h urine collections—**SROC curves were derived from measurements of uMNs in 24-h urine collections and pMNs in blood samples obtained in the supine and the seated position (Fig. 3). The point estimators of sensitivities were significantly different for studies complying with the supine position compared with seated blood sampling conditions (95%, CI 92–97% versus 89%, CI 84–93%; p<0.02), but not for studies using 24-h urine collections (93%, CI 88–95%). In contrast, the point estimators of specificities were significantly different for supine sampling conditions compared with 24-h urine collections (95%, CI 92–97% versus 90%, CI 84–93%; p<0.03), but not for the seated sampling position (94%, CI 90–96%). There was no interaction effect between sample/sampling condition, i.e. blood obtained in a supine or seated position or 24-h urine collections, and analytical method, i.e. IA or HPLC, with only a marginal and statistically not significant increase in likelihood. Values of the partial AUC were 0.942, 0.913, and 0.932 for blood sampling in the supine and the seated position, and for 24-h urine collections, respectively.

### 3.4.    Additional analysis

**3.4.1.    Relation between test accuracy and disease prevalence—**To determine the relationship between test accuracy and disease prevalence according to sample specimen and collecting conditions the equation presented in Methods (see 2.3. Statistical Analysis) was applied using prevalence levels between 0.0 and 1.0 and the individual point estimators of each of the three respective sensitivity/specificity pairs for 24-h urine collections and the supine and seated blood sampling positions. With this approach a steady linear increase in test accuracy from 90% to 93% could be demonstrated for measurements in urine and a steady linear decrease from 94% to 89% for seated blood sampling conditions. In contrast, test accuracy was constant at 95% over the entire range of prevalence levels for the supine sampling position (Fig. 4).

## Discussion

In this meta-analysis we systematically reviewed currently recommended biochemical tests for their discriminative properties and predictive abilities in the diagnosis of P/PGL. As the main result, we found excellent overall diagnostic performance but also distinct differences among the various tests that should be known to users according to clinical application.

First of all, although sensitivities and specificities for measurements of pMNs and uMNs derived from sROC curves of the studies using either IA or HPLC analysis techniques were not significantly different, calculation of the respective partial AUCs revealed a notably lower value for IA methods. Similarly, sensitivities for measurements of pMNs in blood samples obtained in the supine position and of uMNs in 24-h urine collections as well as specificities for the supine and seated sampling positions were not different. However, sensitivity was significantly higher for the supine sampling position compared with seated conditions, and specificity was significantly higher for supine sampling conditions compared with 24-h urine collections. Also, the values of the partial AUCs were clearly different, highest for blood sampling in the supine position, intermediate for measurements in 24-h urine, and lowest for blood sampling under seated conditions. Finally, test accuracy, that is the ratio of correct assessments to all assessments, again was highest for supine blood sampling conditions, and additionally, independent of prevalence levels. In contrast, test accuracy was intermediate and inversely dependent on prevalence levels for the seated sampling position, and it was lowest but directly dependent on prevalence levels for measurements in 24-h urine collections with the two linear plots intersecting at a prevalence level of 50%. Thus, applying three different measures of diagnostic accuracy [22], including one that addressed the issue of prevalence, results either tended to favor or showed a significantly better discriminative potential of HPLC techniques plus blood sampling in the supine position over IA methods, seated blood sampling conditions or measurements in 24-h urine collections, thereby supporting its use as first-line diagnostic test procedure for the biochemical diagnosis of P/PGL [6].

There are however also plausible arguments against this strategy. P/PGL occur with an extremely low annual incidence rate of less than 10 cases per million population [23,24], and the symptomatology of these tumors is protean and often nonspecific [2]. Hypertension is still the most frequent clinical sign encountered in up to 90% of patients [25], whereas the prevalence of P/PGL among adult hypertensive subjects amounts to only 0.2%−0.6% [26–29]. On the other hand, P/PGL may lead to potentially life-threatening cardiovascular and serious other organ complications [30] calling for a timely diagnosis and a low threshold level of clinical suspicion. Consequently, biochemical testing will most probably be carried out in many instances when chances are in fact scant that P/PGL will be detected. However, our results show, that the superior diagnostic accuracy of HPLC plus blood sampling in the supine position does not translate into a relevant gain in diagnostic certainty under these conditions. Indeed, we found that, regardless of the test method used, a negative test result will in all probability correctly predict the absence of P/PGL even assuming prevalence levels of up to tenfold higher than normally encountered in a hypertensive population [31–33], while conversely, the rate of false positive results will be unacceptably high with any one of the tests, and, hence, invariably require confirmatory testing, depending on the extent of elevation (Table 2).

Still another aspect deserves attention. If in fact the vast majority of patients with hypertension initially tested for the presence of P/PGL will correctly be tested negative with any one of the tests currently in use, then considerations such as local preferences, feasibility, and economic efficiency come into play [34]. This holds particularly true for the situation when the anticipated test volume will be too small to counterbalance the substantial

capital expenditures for installation and operation of HPLC as opposed to IA methods [14]. Our results indicate that in this case commitment to the principles of cost-containment does not contradict the general rule that use of a diagnostic procedure should be encouraged if it provides more accurate information and is virtually risk free [35].

Obviously, however, results of between-test comparisons obtained in an average low risk adult hypertensive population cannot simply be extrapolated to the situation in a specialized hypertension/endocrinology referral center. Our analysis highlights the fact, that in this setting use of for instance IA plus measurement of uMNs in 24-h urine collections as the initial diagnostic test procedure will yield a markedly higher rate of false positive results compared to blood sampling in the supine position plus measurement of pMNs with HPLC, and, even more important, the rate of false positive tests will remain considerably higher during any subsequent testing. Whether or not functional tests such as clonidine suppression [36,37] will aid in confirming the diagnosis under these circumstances is unknown. Therefore, proponents of the primary use of IA plus measurement of uMNs in 24-h urine collections for the biochemical diagnosis of P/PGL need to be aware of this potential trade–off and its continued impact on the further diagnostic process.

Our analysis has strengths and limitations. Key strengths are the number and quality of the included studies and the balanced distribution of the different test procedures between these studies. Important limitations are that only 2 studies used age-adjusted diagnostic cutoff levels for pMNs [10,11], and only 3 studies examined the effect of the additional measurement of MTY on test performance [9–11]. Results of the present analysis accordingly allow no conclusions with respect to the diagnostic benefit of these measures. Nevertheless adoption of measurement of MTY and age-adjustment is recommended by the most recent Endocrine Society Clinical Practice Guideline [6], which it is hoped will also help to expedite clarification of this unresolved issue. Another limitation of our analysis concerns the introduction of a continuity correction factor to avoid zero values in the calculation of frequencies. However, this factor was applied to all observed frequencies. Hence, we are confident that any undue effect on our results through the use of this factor can be excluded. Finally, even though the existing evidence for currently reported prevalence rates of P/PGL in the general adult hypertensive population is inconsistent [26–29], we did apply these values, since, if at all, they would obviously represent upper possible limits and therefore not have influenced the overall conclusions of our analysis.

In summary, there is no perfectly precise test for the biochemical diagnosis of P/PGL. The decision about which test to use should therefore be guided primarily by the intentions and expectations of the user. Thus, if exclusion of P/PGL is the overriding goal, for instance IA measurement of uMNs may still be a reasonable choice. If, on the other hand, biochemical proof of P/PGL is the objective, it is highly recommended to immediately proceed with measurement of pMNs in blood samples obtained in the supine position using HPLC with EC or MS/MS detection.

## Acknowledgement

## References

1. Manger WM, Gifford RW: Clinical and experimental pheochromocytoma. Blackwell Science, (1996)

2. Manger WM: The protean manifestations of pheochromocytoma. Horm Metab Res 41(9), 658–663 (2009) [PubMed: 19242899]

3. Lenders JW, Eisenhofer G, Armando I, Keiser HR, Goldstein DS, Kopin IJ: Determination of metanephrines in plasma by liquid chromatography with electrochemical detection. Clin Chem 39(1), 97–103 (1993) [PubMed: 8419068]

4. Lenders JW, Pacak K, Walther MM, Linehan WM, Mannelli M, Friberg P, Keiser HR, Goldstein DS, Eisenhofer G: Biochemical diagnosis of pheochromocytoma: which test is best? Jama 287(11), 1427–1434 (2002) [PubMed: 11903030]

5. Davidson DF: Phaeochromocytoma with normal urinary catecholamines: the potential value of urinary free metadrenalines. Ann Clin Biochem 39(Pt 6), 557–566 (2002) [PubMed: 12564837]

6. Lenders JW, Duh QY, Eisenhofer G, Gimenez-Roqueplo AP, Grebe SK, Murad MH, Naruse M, Pacak K, Young WF, Jr., Endocrine S: Pheochromocytoma and paraganglioma: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 99(6), 1915–1942 (2014) [PubMed: 24893135]

7. Eisenhofer G: Free or total metanephrines for diagnosis of pheochromocytoma: what is the difference? Clin Chem 47(6), 988–989 (2001) [PubMed: 11375282]

8. Eisenhofer G, Lenders JW, Timmers H, Mannelli M, Grebe SK, Hofbauer LC, Bornstein SR, Tiebel O, Adams K, Bratslavsky G, Linehan WM, Pacak K: Measurements of plasma methoxytyramine, normetanephrine, and metanephrine as discriminators of different hereditary forms of pheochromocytoma. Clin Chem 57(3), 411–420 (2011) [PubMed: 21262951]

9. Brain KL, Kay J, Shine B: Measurement of urinary metanephrines to screen for pheochromocytoma in an unselected hospital referral population. Clin Chem 52(11), 2060–2064 (2006) [PubMed: 16990424]

10. Därr R, Pamporaki C, Peitzsch M, Miehle K, Prejbisz A, Peczkowska M, Weismann D, Beuschlein F, Sinnott R, Bornstein SR, Neumann HP, Januszewicz A, Lenders J, Eisenhofer G: Biochemical diagnosis of phaeochromocytoma using plasma-free normetanephrine, metanephrine and methoxytyramine: importance of supine sampling under fasting conditions. Clin Endocrinol (Oxf) 80(4), 478–486 (2014) [PubMed: 24102244]

11. Weismann D, Peitzsch M, Raida A, Prejbisz A, Gosk M, Riester A, Willenberg HS, Klemm R, Manz G, Deutschbein T, Kroiss M, Därr R, Bidlingmaier M, Januszewicz A, Eisenhofer G, Fassnacht M: Measurements of plasma metanephrines by immunoassay vs liquid chromatography with tandem mass spectrometry for diagnosis of pheochromocytoma. Eur J Endocrinol 172(3), 251–260 (2015) [PubMed: 25452465]

12. Pillai D, Ross HA, Kratzsch J, Pedrosa W, Kema I, Hoad K, Rouaix N, Fathi M, Nader H, Mathian B, Grouzmann E: Proficiency test of plasma free and total metanephrines: report from a study group. Clin Chem Lab Med 47(6), 786–790 (2009) [PubMed: 19445649]

13. Pillai D, Callen S: Pilot quality assurance programme for plasma metanephrines. Ann Clin Biochem 47(Pt 2), 137–142 (2010) [PubMed: 20144968]

14. Eisenhofer G, Peitzsch M, McWhinney BC: Impact of LC-MS/MS on the laboratory diagnosis of catecholamine-producing tumors. Trends Analyt Chem 84(B), 106–116 (2016)

15. de Jong WH, Eisenhofer G, Post WJ, Muskiet FA, de Vries EG, Kema IP: Dietary influences on plasma and urinary metanephrines: implications for diagnosis of catecholamine-producing tumors. J Clin Endocrinol Metab 94(8), 2841–2849 (2009) [PubMed: 19567530]

16. Lenders JW, Willemsen JJ, Eisenhofer G, Ross HA, Pacak K, Timmers HJ, Sweep CG: Is supine rest necessary before blood sampling for plasma metanephrines? Clin Chem 53(2), 352–354 (2007) [PubMed: 17200132]

17. Sawka AM, Prebtani AP, Thabane L, Gafni A, Levine M, Young WF, Jr.: A systematic review of the literature examining the diagnostic efficacy of measurement of fractionated plasma free metanephrines in the biochemical diagnosis of pheochromocytoma. BMC Endocr Disord 4(1), 2 (2004) [PubMed: 15225350]

18. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P: Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 151(4), 264–269, W264 (2009) [PubMed: 19622511]

19. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG, Standards for Reporting of Diagnostic, A.: The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Ann Intern Med 138(1), W1–12 (2003) [PubMed: 12513067]

20. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG, Standards for Reporting of Diagnostic, A.: The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem 49(1), 7–18 (2003) [PubMed: 12507954]

21. Reitsma J, Glas A, Rutjes A, Scholten R, Bossuyt P, Zwinderman A: Bivariate Analysis of Sensitivity and Specificity Produces Informative Summary Measures in Diagnostic Reviews. J Clin Epidemiol 58(10), 982–990 (2005) [PubMed: 16168343]

22. Mallett S, Halligan S, Thompson M, Collins GS, Altman DG: Interpreting diagnostic accuracy studies for patient care. BMJ 345, e3999 (2012) [PubMed: 22750423]

23. Beard CM, Sheps SG, Kurland LT, Carney JA, Lie JT: Occurrence of pheochromocytoma in Rochester, Minnesota, 1950 through 1979. Mayo Clin Proc 58(12), 802–804 (1983) [PubMed: 6645626]

24. Stenström G, Svärdsudd K: Pheochromocytoma in Sweden 1958–1981. An analysis of the National Cancer Registry Data. Acta Med Scand 220(3), 225–232 (1986) [PubMed: 3776697]

25. Zuber SM, Kantorovich V, Pacak K: Hypertension in pheochromocytoma: characteristics and treatment. Endocrinol Metab Clin North Am 40(2), 295–311, vii (2011) [PubMed: 21565668]

26. Omura M, Saito J, Yamaguchi K, Kakuta Y, Nishikawa T: Prospective study on the prevalence of secondary hypertension among hypertensive patients visiting a general outpatient clinic in Japan. Hypertens Res 27(3), 193–202 (2004) [PubMed: 15080378]

27. Ariton M, Juan CS, AvRuskin TW: Pheochromocytoma: clinical observations from a Brooklyn tertiary hospital. Endocr Pract 6(3), 249–252 (2000) [PubMed: 11421540]

28. Anderson GH, Jr., Blakeman N, Streeten DH: The effect of age on prevalence of secondary forms of hypertension in 4429 consecutively referred patients. J Hypertens 12(5), 609–615 (1994) [PubMed: 7930562]

29. Sinclair AM, Isles CG, Brown I, Cameron H, Murray GD, Robertson JW: Secondary hypertension in a blood pressure clinic. Arch Intern Med 147(7), 1289–1293 (1987) [PubMed: 3606286]

30. Brouwers FM, Lenders JW, Eisenhofer G, Pacak K: Pheochromocytoma as an endocrine emergency. Rev Endocr Metab Disord 4(2), 121–128 (2003) [PubMed: 12766539]

31. Mantero F, Terzolo M, Arnaldi G, Osella G, Masini AM, Ali A, Giovagnetti M, Opocher G, Angeli A: A survey on adrenal incidentaloma in Italy. Study Group on Adrenal Tumors of the Italian Society of Endocrinology. J Clin Endocrinol Metab 85(2), 637–644 (2000) [PubMed: 10690869]

32. Mansmann G, Lau J, Balk E, Rothberg M, Miyachi Y, Bornstein SR: The clinically inapparent adrenal mass: update in diagnosis and management. Endocr Rev 25(2), 309–340 (2004) [PubMed: 15082524]

33. Därr R, Lenders JW, Hofbauer LC, Naumann B, Bornstein SR, Eisenhofer G: Pheochromocytoma - update on disease management. Ther Adv Endocrinol Metab 3(1), 11–26 (2012) [PubMed: 23148191]

34. Sawka AM, Gafni A, Thabane L, Young WF, Jr.: The economic implications of three biochemical screening algorithms for pheochromocytoma. J Clin Endocrinol Metab 89(6), 2859–2866 (2004) [PubMed: 15181069]

35. Pauker SG, Kassirer JP: The threshold approach to clinical decision making. N Engl J Med 302(20), 1109–1117 (1980) [PubMed: 7366635]

36. Därr R, Lenders JW, Stange K, Kindel B, Hofbauer LC, Bornstein SR, Eisenhofer G: [Diagnosis of pheochromocytoma and paraganglioma: the clonidine suppression test in patients with borderline elevations of plasma free normetanephrine]. Dtsch Med Wochenschr 138(3), 76–81 (2013) [PubMed: 23299341]

37. Eisenhofer G, Goldstein DS, Walther MM, Friberg P, Lenders JW, Keiser HR, Pacak K: Biochemical diagnosis of pheochromocytoma: how to distinguish true- from false-positive test results. J Clin Endocrinol Metab 88(6), 2656–2666 (2003) [PubMed: 12788870]

38. Raber W, Raffesberg W, Bischof M, Scheuba C, Niederle B, Gasic S, Waldhausl W, Roden M: Diagnostic efficacy of unconjugated plasma metanephrines for the detection of pheochromocytoma. Arch Intern Med 160(19), 2957–2963 (2000) [PubMed: 11041903]

39. Sawka AM, Jaeschke R, Singh RJ, Young WF, Jr.: A comparison of biochemical tests for pheochromocytoma: measurement of fractionated plasma metanephrines compared with the combination of 24-hour urinary metanephrines and catecholamines. J Clin Endocrinol Metab 88(2), 553–558 (2003) [PubMed: 12574179]

40. Giovanella L, Squin N, Ghelfo A, Ceriani L: Chromogranin A immunoradiometric assay in diagnosis of pheochromocytoma: comparison with plasma metanephrines and 123I-MIBG scan. Q J Nucl Med Mol Imaging 50(4), 344–347 (2006) [PubMed: 17043632]

41. Boyle JG, Davidson DF, Perry CG, Connell JM: Comparison of diagnostic accuracy of urinary free metanephrines, vanillyl mandelic Acid, and catecholamines and plasma catecholamines for diagnosis of pheochromocytoma. J Clin Endocrinol Metab 92(12), 4602–4608 (2007) [PubMed: 17635948]

42. Perry CG, Sawka AM, Singh R, Thabane L, Bajnarek J, Young WF, Jr.: The diagnostic efficacy of urinary fractionated metanephrines measured by tandem mass spectrometry in detection of pheochromocytoma. Clin Endocrinol (Oxf) 66(5), 703–708 (2007) [PubMed: 17388796]

43. Vaclavik J, Stejskal D, Lacnak B, Lazarova M, Jedelsky L, Kadalova L, Janosova M, Frysak Z, Vlcek P: Free plasma metanephrines as a screening test for pheochromocytoma in low-risk patients. J Hypertens 25(7), 1427–1431 (2007) [PubMed: 17563565]

44. Gao YC, Lu HK, Luo QY, Chen LB, Ding Y, Zhu RS: Comparison of free plasma metanephrines enzyme immunoassay with (131)I-MIBG scan in diagnosis of pheochromocytoma. Clin Exp Med 8(2), 87–91 (2008) [PubMed: 18618218]

45. Hickman PE, Leong M, Chang J, Wilson SR, McWhinney B: Plasma free metanephrines are superior to urine and plasma catecholamines and urine catecholamine metabolites for the investigation of phaeochromocytoma. Pathology 41(2), 173–177 (2009) [PubMed: 19152190]

46. Procopiou M, Finney H, Akker SA, Chew SL, Drake WM, Burrin J, Grossman AB: Evaluation of an enzyme immunoassay for plasma-free metanephrines in the diagnosis of catecholamine-secreting tumors. Eur J Endocrinol 161(1), 131–140 (2009) [PubMed: 19383804]

47. Grouzmann E, Drouard-Troalen L, Baudin E, Plouin PF, Muller B, Grand D, Buclin T: Diagnostic accuracy of free and total metanephrines in plasma and fractionated metanephrines in urine of patients with pheochromocytoma. Eur J Endocrinol 162(5), 951–960 (2010) [PubMed: 20142367]

48. Peaston RT, Graham KS, Chambers E, van der Molen JC, Ball S: Performance of plasma free metanephrines measured by liquid chromatography-tandem mass spectrometry in the diagnosis of pheochromocytoma. Clin Chim Acta 411(7–8), 546–552 (2010) [PubMed: 20083099]

49. Christensen TT, Frystyk J, Poulsen PL: Comparison of plasma metanephrines measured by a commercial immunoassay and urinary catecholamines in the diagnosis of pheochromocytoma. Scand J Clin Lab Invest 71(8), 695–700 (2011) [PubMed: 22023043]

50. Sarathi V, Pandit R, Jagtap V, Lila AR, Bandgar TR, Menon PS, Varthakavi P, Raghavan VP, Shah NS: Performance of plasma fractionated free metanephrines by enzyme immunoassay in the diagnosis of pheochromocytoma and paraganglioma. Endocr Pract 17(5), 759–765 (2011) [PubMed: 21742610]

51. Mullins F, O'Shea P, FitzGerald R, Tormey W: Enzyme-linked immunoassay for plasma-free metanephrines in the biochemical diagnosis of phaeochromocytoma in adults is not ideal. Clin Chem Lab Med 50(1), 105–110 (2011) [PubMed: 21981378]

52. Unger N, Hinrichs J, Deutschbein T, Schmidt H, Walz MK, Mann K, Petersenn S: Plasma and urinary metanephrines determined by an enzyme immunoassay, but not serum chromogranin A for the diagnosis of pheochromocytoma in patients with adrenal mass. Exp Clin Endocrinol Diabetes 120(8), 494–500 (2012) [PubMed: 22696168]

53. Jeyaraman K, Natarajan V, Thomas N, Jacob PM, Nair A, Shanthly N, Oommen R, Varghese G, Joseph FJ, Seshadri MS, Rajaratnam S: The role of urinary fractionated metanephrines in the

diagnosis of phaeochromocytoma. Indian J Med Res 137(2), 316–323 (2013) [PubMed: 23563375]

54. Pussard E, Chaouch A, Said T: Radioimmunoassay of free plasma metanephrines for the diagnosis of catecholamine-producing tumors. Clin Chem Lab Med 52(3), 437–444 (2014) [PubMed: 24088613]

55. Tanaka Y, Isobe K, Ma E, Imai T, Kikumori T, Matsuda T, Maeda Y, Sakurai A, Midorikawa S, Hataya Y, Kato T, Kamide K, Ikeda Y, Okada Y, Adachi M, Yanase T, Takahashi H, Yokoyama C, Arai Y, Hashimoto K, Shimano H, Hara H, Kawakami Y, Takekoshi K: Plasma free metanephrines in the diagnosis of pheochromocytoma: diagnostic accuracy and strategies for Japanese patients. Endocr J 61(7), 667–673 (2014) [PubMed: 24871964]

56. Kim HJ, Lee JI, Cho YY, Lee SY, Kim JH, Jung BC, Kim SW, Chung JH, Min YK, Lee MS, Lee MK, Kim JH: Diagnostic accuracy of plasma free metanephrines in a seated position compared with 24-hour urinary metanephrines in the investigation of pheochromocytoma. Endocr J 62(3), 243–250 (2015) [PubMed: 25476661]
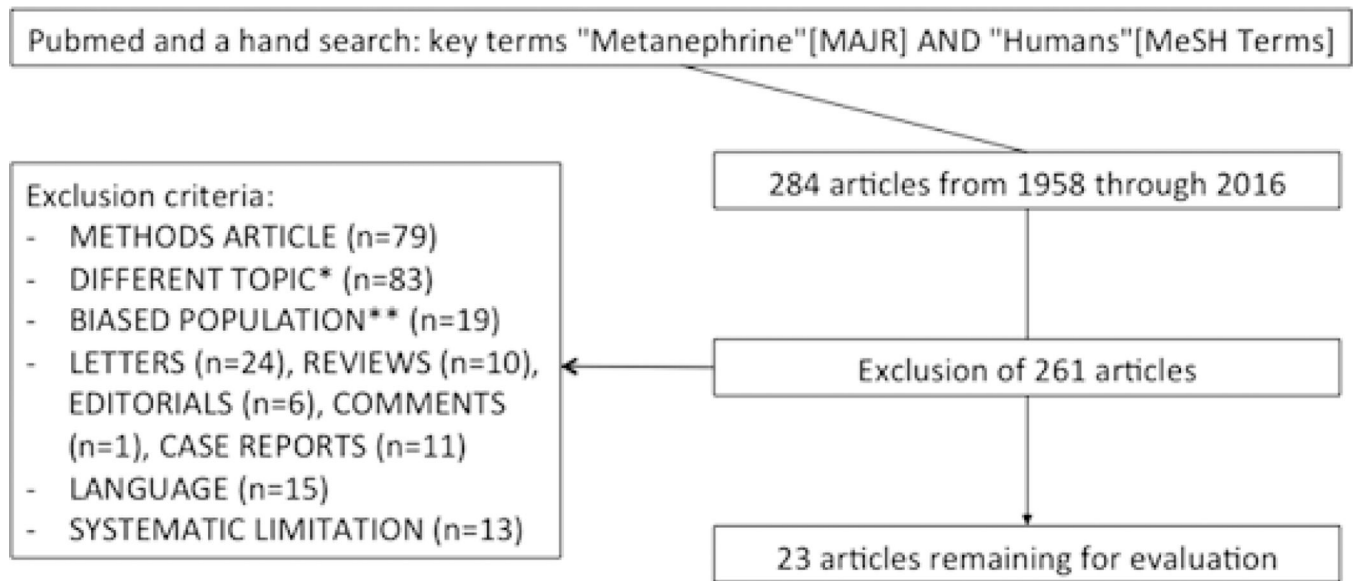
**Fig. 1.**
Schematic representation of the study selection process. *e.g. physiology; neurology; psychiatry; obstetrics; surgery; periodontology; forensic medicine; diagnostic radiology; confounding variables; differential diagnosis; tumor type, size, recurrence and follow up; analytes; sampling periods; repeated and dynamic testing; metabolism; economic considerations. **e.g. lack of control group; healthy volunteers; patients with Head & Neck or hereditary P/PGL; uni- or bilaterally adrenalectomized patients; children

**Fig. 2.**
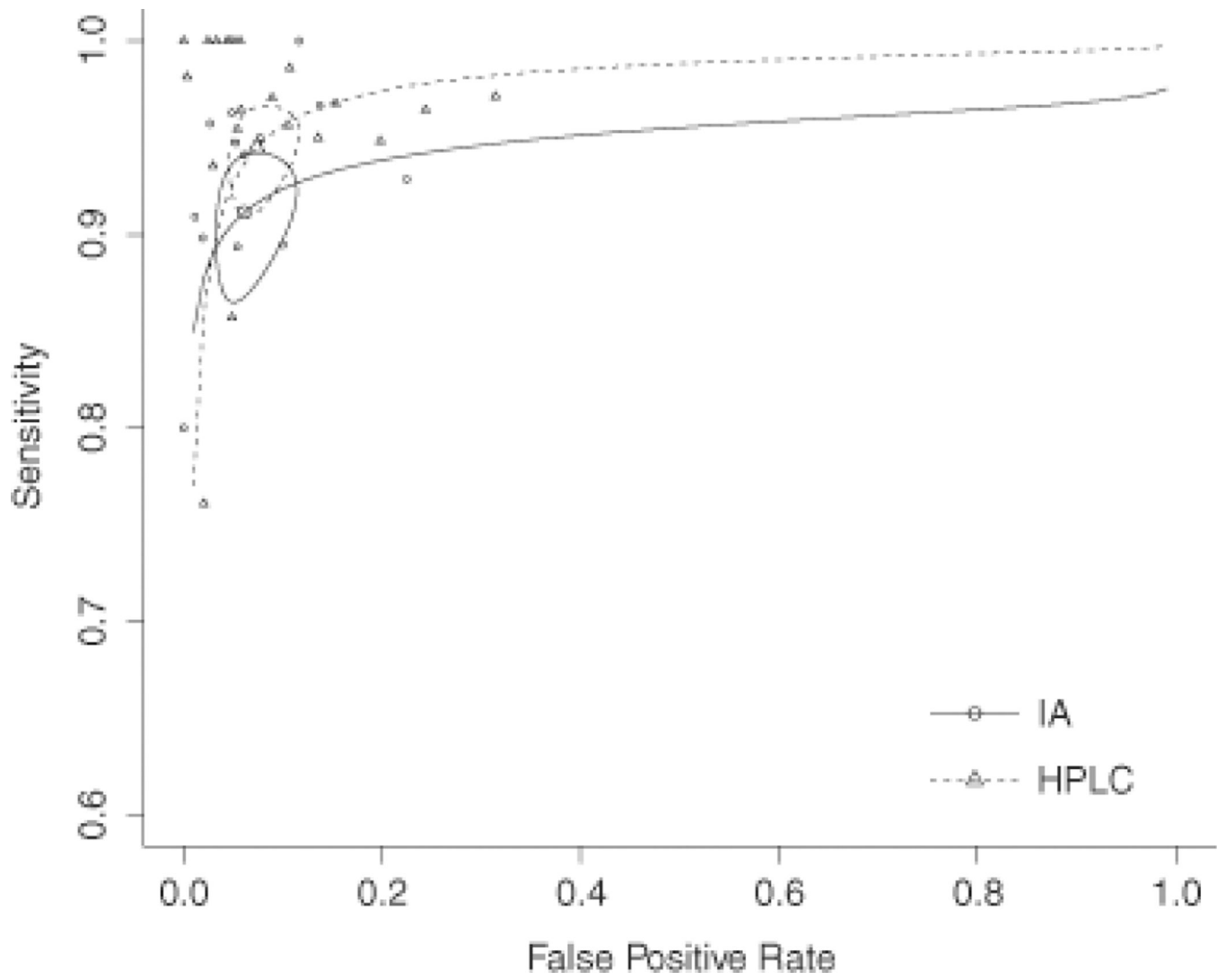sROC curves for immunoassay (IA) and high performance liquid chromatography (HPLC) analysis techniques. Circles represent confidence intervals
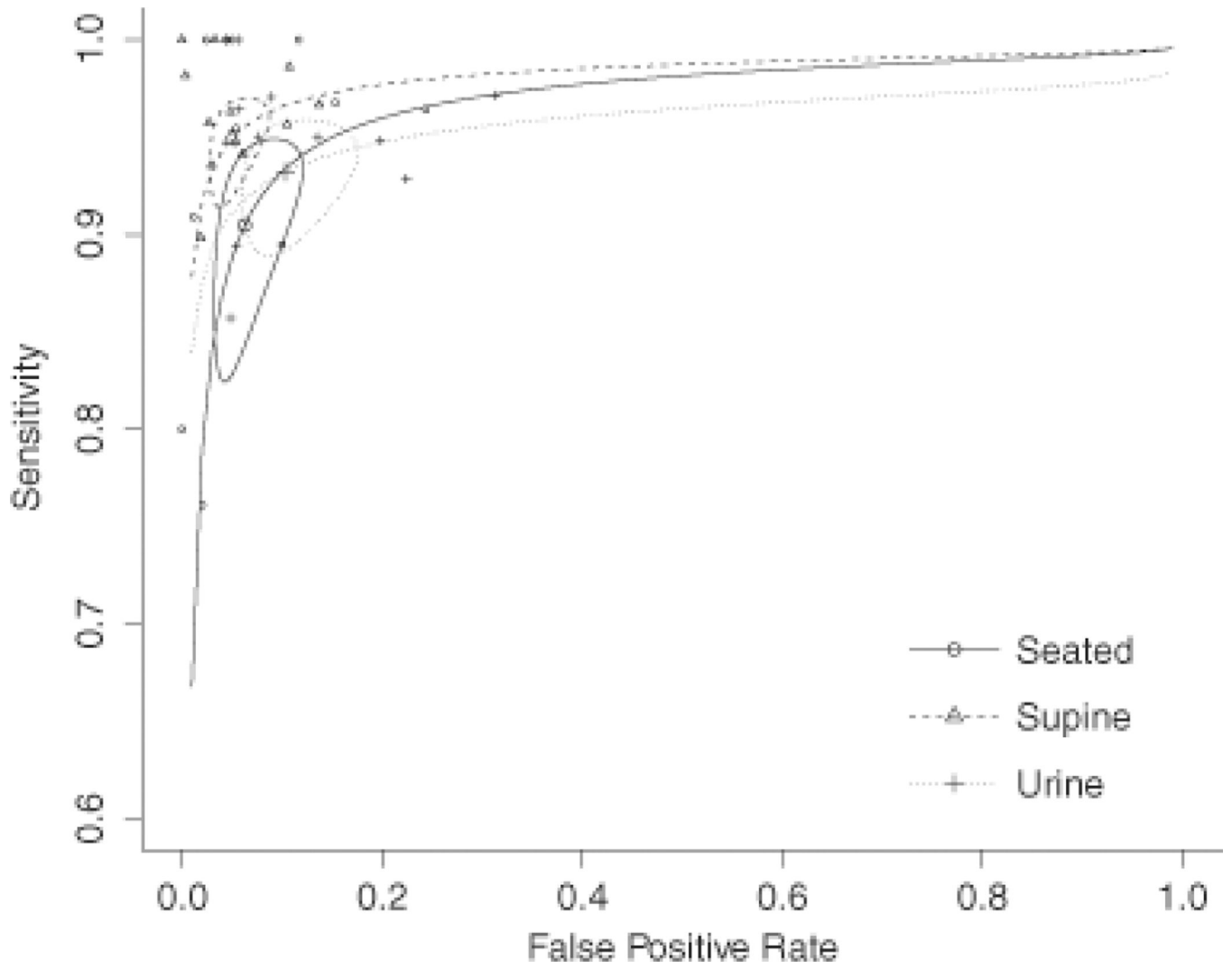
**Fig. 3.**
sROC curves for seated and supine blood sampling conditions and 24-h urine collections.
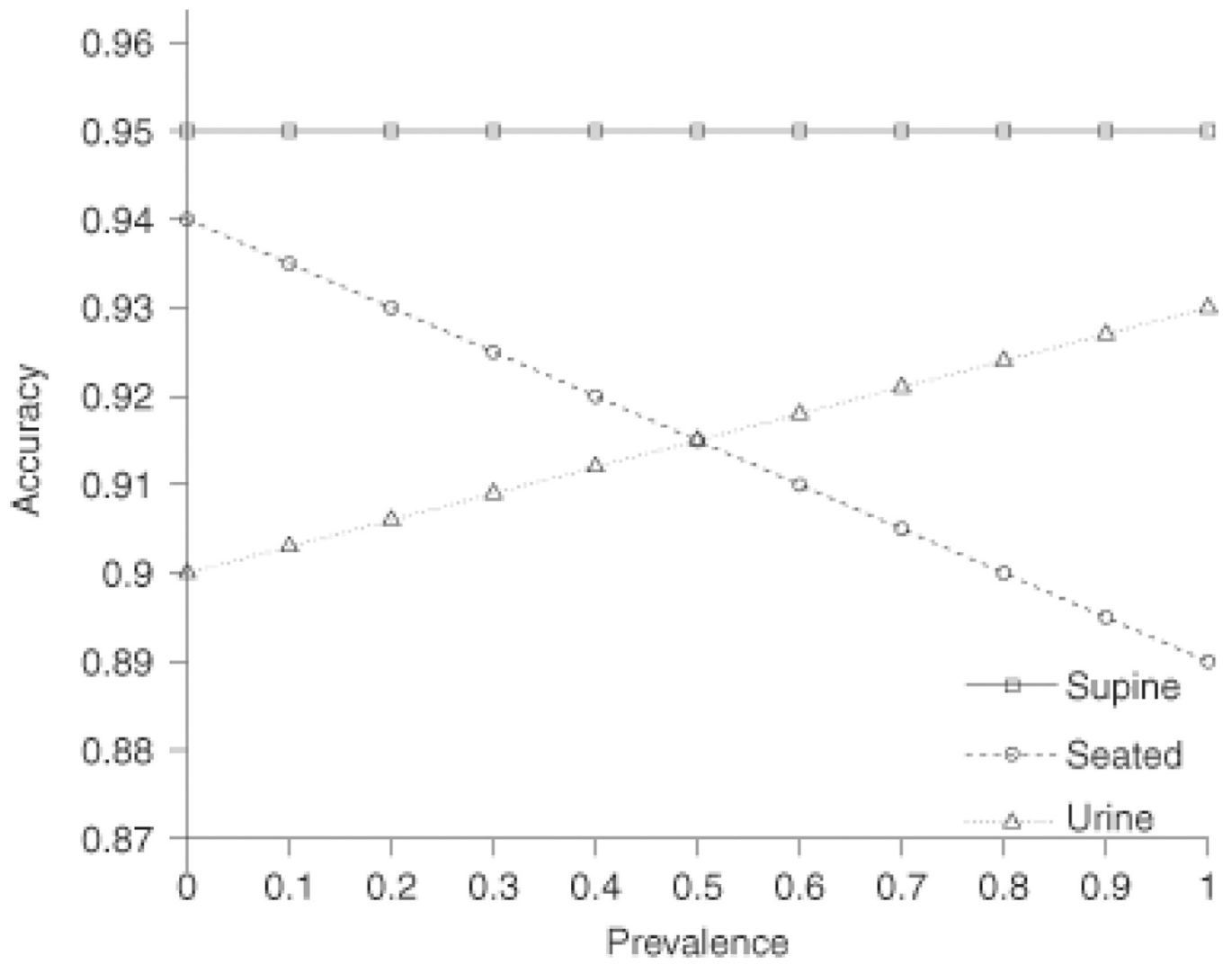Circles represent confidence intervals

**Fig. 4.**
Test performance according to disease prevalence for blood sampling in a seated or supine position and for 24-h urine collections

**Table 1.**

Diagnostic performance of measurements of uMNs in 24-h urine collections and pMNs from blood obtained in a supine or seated position of included studies using HPLC or IA analysis techniques. Year of publication, total number of subjects per study and reported sensitivities and specificities are indicated. Figures in parentheses represent the rate of correct assessments.

| No. | Study | Year | Sample | Method | Sensitivity | Specificity | Total Subjects |
|---|---|---|---|---|---|---|---|
| 1 | Raber et al [38] | 2000 | Supine | HPLC-EC | 100% (17/17) | 100% (14/14) | 31 |
| 2 | Lenders et al [4] | 2002 | Supine | HPLC-EC | 99% (211/214) | 89% (575/644) | 858 |
| | | | Urine | HPLC-EC | 97% (102/105) | 69% (310/452) | 557 |
| 3 | Sawka et al [39] | 2003 | Seated | HPLC-EC | 97% (30/31) | 85% (221/261) | 292 |
| 4 | Brain et al [9] | 2006 | Urine | HPLC-CD | 100% (14/14) | 96% (1724/1805) | 1819 |
| 5 | Giovanella et al [40] | 2006 | Seated [a] | HPLC-EC | 95% (42/44) | 95% (140/148) | 192 |
| 6 | Boyle et al [41] | 2007 | Urine [f] | HPLC-CD | 100% (25/25) | 94% (116/123) | 148 |
| 7 | Perry et al [42] | 2007 | Urine | HPLC-MS/MS | 97% (99/102) | 91% (368/404) | 506 |
| 8 | Vaclavik et al [43] | 2007 | Supine | HPLC-EC | 100% (25/25) | 97% (1194/1235) | 1260 |
| 9 | Gao et al [44] | 2008 | Supine | EIA | 97% (29/30) | 86% (44/51) | 81 |
| 10 | Hickman et al [45] | 2009 | Seated [a] | HPLC-EC | 100% (14/14) [c] | 98% (40/41) | 55 |
| | | | Urine | HPLC-EC | 86% (12/14) [c] | 95% (39/41) | 55 |
| 11 | Procopiou et al [46] | 2009 | Seated [a] | EIA | 80% (16/20) [b] | 100% (156/156) | 176 |
| 12 | Grouzmann et al [47] | 2010 | Supine | HPLC-EC | 96% (44/46) | 89% (102/114) | 160 |
| | | | Urine | HPLC-AC | 95% (38/40) | 86% (121/140) | 180 |
| 13 | Peaston et al [48] | 2010 | Seated | HPLC-MS/MS | 100% (38/38) | 96% (108/113) | 151 |
| | | | Seated | ELISA | 95% (36/38) | 95% (107/113) | 151 |
| 14 | Christensen et al [49] | 2011 | Seated [e] | EIA | 91% (10/11) | 99% (172/174) | 185 |
| 15 | Sarathi et al [50] | 2011 | Supine [c] | EIA | 94% (32/34) | 94% (62/66) | 100 |
| 16 | Mullins et al [51] | 2012 | Seated | ELISA | 100% (13/13) | 88% (53/60) [c] | 73 |
| | | | Seated | HPLC-MS/MS | 100% (13/13) | 95% (57/60) | 73 |
| 17 | Unger et al [52] | 2012 | Seated | EIA | 89% (17/19) | 90% (54/60) | 79 |
| | | | Urine | EIA | 93% (13/14) | 78% (38/49) | 63 |
| 18 | Jeyaraman et al [53] | 2013 | Urine | ELISA | 95% (19/20) | 92% (48/52) | 72 |

| No. | Study | Year | Sample | Method | Sensitivity | Specificity | Total Subjects |
|---|---|---|---|---|---|---|---|
| 19 | Daerr et al [10] [d] | 2014 | Seated | HPLC-MS/MS | 76% (51/67) | 98% (191/195) | 262 |
| | | 2014 | Supine | HPLC-MS/MS | 94% (58/62) | 97% (425/438) | 500 |
| | | 2014 | Urine | HPLC-MS/MS | 95% (55/58) | 80% (352/439) | 497 |
| 20 | Pussard et al [54] | 2014 | Supine | RIA | 90% (53/59) | 98% (465/474) | 533 |
| 21 | Tanaka et al [55] | 2014 | Supine [a] | EIA | 96% (45/47) | 97% (36/37) | 84 |
| | | 2014 | Urine | EIA | 89% (42/47) | 95% (35/37) | 84 |
| 22 | Kim et al [56] | 2015 | Seated | HPLC-MS/MS | 96% (27/28) | 76% (118/156) | 184 |
| | | 2015 | Urine | HPLC-EC | 96% (27/28) | 94% (147/156) | 184 |
| 23 | Weismann et al [11] | 2015 | Supine | HPLC-MS/MS | 98% (53/54) | 100% (286/287) | 341 |
| | | 2015 | Supine | EIA | 96% (52/54) | 95% (273/287) | 341 |

*HPLC-CD, HPLC-EC* and *HPLC-MS/MS* high performance liquid chromatography with coulometric, electrochemical and tandem mass spectrometric detection, respectively; *ELISA* enzyme-linked immunoassay, *EIA* enzyme immunoassay, *RIA* radioimmunoassay

[a] personal communications

[b] Pheochromocytoma (P) and paraganglioma (PGL) were considered in the analysis. Two dopamine producing PGLs with negative nor- and metanephrine (NMN and MN) were excluded, methoxytyramine (MTY) levels were not measured

[c] Discordant to the Endocrine Society Practice Guidelines (Table 4)[6], blood samples were drawn in the supine position for Sarathi *et al* [50], numbers were restricted to 14 patients with pheochromocytoma for Hickman *et al* [45], and 7 patients had false positive test results in the immunoassay for Mullins *et al* [51]

[d] Data of this study were reanalyzed using only the results for NMN and MN in urine and plasma, while MTY results were omitted to ensure comparability with the various studies

[e] one biochemical silent tumor was classified as ganglioneuroma

[f] urinary free fractionated MNs were measured [5]

**Table 2.**

Diagnostic accuracy of biochemical tests in the diagnosis of P/PGL

| | Studies [number] | SE [%] | SP [%] | AUC [partial][a] | TA [%] | NPV [0.5%][b] | [5%][c] | PPV [0.5%][b] | [5%][c] |
|---|---|---|---|---|---|---|---|---|---|
| **HPLC** | 15 | 94 | 93 | 0.947 | 93–94 | n.a. | | n.a. | |
| **IA** | 11 | 91 | 93 | 0.911 | 91–93 | n.a. | | n.a. | |
| **Supine** | 10 | 95[*] | 95[**] | 0.942 | 95–95 | 0.9997 | 0.9972 | 0.0872 | 0.5000 |
| **Seated** | 10 | 89[*] | 94 | 0.913 | 89–94 | 0.9994 | 0.9939 | 0.0694 | 0.4384 |
| **Urine** | 11 | 93 | 90[**] | 0.932 | 90–93 | 0.9996 | 0.9959 | 0.0446 | 0.3286 |

*SE* sensitivity; *SP* specificity; *TA* test accuracy range (for prevalence ranges between 0.0–1.0); *NPV* negative predictive value; *PPV* positive predictive value; *n.a.* not assessed

[*]
supine versus seated p<0.02;

[**]
supine versus 24-h urine p<0.03

[a]
for details see Methods

[b]
prevalence of P/PGL among hypertensive subjects (0.2–0.6%)

[c]
prevalence of P/PGL in subjects with incidentaloma

*Endocrine.* Author manuscript; available in PMC 2018 December 18.