

Published in final edited form as:

*Clin Chem.* 2006 November ; 52(11): 2060–2064. doi:10.1373/clinchem.2006.070805.

## Measurement of urinary metanephrines to screen for pheochromocytoma in an unselected hospital referral population

Keith L. Brain<sup>1</sup>, Jonathan Kay<sup>2</sup>, and Brian Shine<sup>2</sup>

<sup>1</sup>Department of Pharmacology, University of Oxford, Mansfield Rd., Oxford OX1 3QT, UK

<sup>2</sup>Department of Clinical Biochemistry, John Radcliffe Hospital, Headley Way, Oxford, OX3 9DU, UK

### Abstract

**Background**—Despite the rarity of pheochromocytoma, the dangers of uncontrolled severe hypertension and the very effective surgical treatment of this condition mean that diagnosis is important. Urinary or plasma catecholamines or catecholamine-derivatives are commonly used to screen for pheochromocytomas prior to imaging. This study investigates whether derived measures obtained from 24-hour urinary metanephrine results, patient age and sex can better predict tumors in populations with a low pre-test probability.

**Methods**—This study takes a pragmatic approach by retrospectively studying the outcomes of an unselected population referred for urinary metanephrine testing (1819 patients) to a tertiary hospital laboratory, and investigates the usefulness of some simple derivative measures for detecting pheochromocytoma. Urinary 24-hour excretion of metanephrine, normetanephrine and 3-methoxytyramine were normalized by dividing by an age- and sex- specific reference range. The ability of products of these normalized measures to predict pheochromocytomas was assessed, compared to a gold standard of biopsy-confirmed tumor.

**Results**—The normalized product of urinary metanephrine and normetanephrine excretion (nMAD.nNMT) proved to be a highly sensitive (100%) and specific (99.1%) measure yielding a positive predictive value 82%. Receiver-operator characteristic curves were not improved by including the normalized 3-methoxytyramine concentrations in the product. nMAD.nNMT gave higher sensitivity and specificity than either test alone.

**Conclusion**—We suggest that nMAD.nNMT is a useful measure for identifying pheochromocytoma in a population with a low pre-test probability.

### Keywords

Pheochromocytoma; urinary metanephrines

### Introduction

The best available screening tests for pheochromocytoma have good sensitivity, yet the rarity of pheochromocytoma in people with suggestive symptoms (such as hypertension) makes the pre-test probability low, and hence the positive predictive value is low even for the best available tests. The most common biochemical approaches used to detect pheochromocytoma are measurements of metanephrines or catecholamines in either plasma

or urine, although some studies have advocated the use of chromogranin (1) or platelet catecholamines (2).

Recent work has advocated the measurement of plasma metanephrines (3), and this position is supported in relevant reviews (4, 5). However, Sawka and colleagues (6) reported that in a screening population, the probability likelihoods for raised plasma metanephrines and urinary total metanephrines were 6.3 and 58.9 respectively. This high probability likelihood for urinary metanephrines implies that it is a more suitable screening test for those with a low pre-test probability, although it might be argued that the high upper limits for detection applied in their study resulted in unacceptable sensitivity. This controversial issue is perhaps best resolved by comparing receiver-operator characteristic curves when evaluating which test might be most useful for screening purposes.

It has been widely accepted that fractionated urinary metanephrines are preferable to the measurement of total urinary metanephrines because they allow detection of tumors that predominantly produce one of the three O-methylated metabolites (7, 8). A common approach to using fractionated urinary metanephrines is to declare a result positive if either normetanephrine or metanephrine is high; for example, see (6). While the sensitivity of urinary fractionated metanephrines for detection of pheochromocytoma is high, specificity is low, at only 45% in patients tested because of suspicion of sporadic forms of the tumor (3). This low specificity creates a need for better interpretation of urinary fractionated metanephrines. In the present study, the outcomes of a cohort of patients referred to a tertiary hospital laboratory for pheochromocytoma screening measuring fractionated urinary metanephrines, over a period of 28 months, are investigated and the utility of derived measures of urinary fractionated metanephrines are compared.

## Materials and methods

A retrospective analysis of all adult 24-hour urinary metanephrine tests referred to the John Radcliffe (Oxford, UK) between 1<sup>st</sup> January 2003 and 4<sup>th</sup> May 2005 was performed. Such urinary metanephrine tests are the current preferred testing strategy of this referral centre for the biochemical diagnosis of pheochromocytoma. Samples from each acidified 24-hour specimen were hydrolyzed in boiling acidic conditions to release conjugated metanephrines. Ion-exchange chromatography, using first a cationic and then an anionic exchange column, preceded isocratic reverse phase HPLC using Bio-Rad Analytical HPLC Cartridges (catalogue number 195-6088; Bio-Rad, Hemel Hempstead, England). Electrochemical detection was with an ESA COULOCHEM II coulometric electrochemical detector (ESA Analytical, Aylesbury, England; first electrode at -45 mV, the second at -350 mV). Regular control solutions were Bio-Rad Urine Standard and Lyphocheck Urine controls (catalogue numbers C-390-10 and C-395-10). The performance of the method was assessed in the UK NEQAS (<http://www.ukneqas.org.uk/>) scheme for urinary metanephrines.

For all specimens in which either normetanephrine, metanephrine or 3-methoxytyramine were above the age- and sex- specific references ranges, and where the patient was seen at the John Radcliffe or a general practice within the referral area, a review of all biochemistry, hematology, immunology and online radiology results was carried out (including an assessment of clinical information provided at the time of the request). In the case of abnormal test results, where the diagnosis was unclear the hospital notes were reviewed.

Each measured value of normetanephrine (NMT) and metanephrine (MAD) was normalized by dividing by the age- and sex- dependent upper limits of normal (9), in the first test after 1<sup>st</sup> January 2003 for each patient. The laboratory references ranges were:

- Metanephrine ( $\mu\text{mol}/24$  hours): male, 1.9; female, 1.4

- Normetanephrine ( $\mu\text{mol}/24$  hours): age 18-40 years, male, 3.6; female, 3.0; age 40-60 years, male 4.25; female 3.45; age greater than 60 years, male, 4.5; female 3.65;
- 3-methoxytyramine ( $\mu\text{mol}/24$  hours): age 18-40 years, 2.75; age 40-60 years, 2.55; age greater than 60 years, 2.3.

The John Radcliffe is a regional referral centre for the treatment of pheochromocytomas, and so would expect to report on most pheochromocytomas occurring in the laboratory referral area. Histopathology reports were retrospectively searched for the terms “phaeochromocytoma”, “pheochromocytoma”, “paraganglioma” or “paraganglionoma”, in order to identify all pheochromocytomas histologically diagnosed in the referral population.

All searches were performed with in-house laboratory database software. Receiver-operator characteristic curves (ROCs) were plotted with Excel for Mac, Version 11 (Microsoft); areas under the curve of ROCs were calculated numerically using Excel. The standard error for calculating the area under ROC curves was calculated with GraphPad Prism (v. 4.0b; GraphPad Software, San Diego, USA), using a method described by Hanley and McNeil (10). When comparing tests, a  $z$  value was calculated where

$$z = \frac{|\text{Area}_1 - \text{Area}_2|}{\sqrt{S.E._1^2 + S.E._2^2 - 2rS.E._1.S.E._2}}$$

where  $\text{Area}_1$  and  $\text{Area}_2$  are the areas being compared,  $S.E._1$  and  $S.E._2$  are their corresponding standard errors and  $r$  is the estimated correlation between the areas, calculated according to (10) with test correlation coefficients calculated with the nonparametric Spearman correlation coefficient. As multiple measures based on the same underlying data are considered, these measures are likely to be correlated. One-tailed  $P$  values were obtained by comparing this  $z$  value with a standard normal distribution.

## Results

In the referral population, 53% were women, the population mean age was 51 years, 23% of tests were requested by local general practitioners (GPs), 21% were from hospital outpatient clinics, and 55% were from hospital wards (or through referral hospital laboratories).

There were 2470 24-hour urinary metanephrine tests from 1819 people performed in the study period, from which there were 149 abnormal tests (either of high metanephrine or normetanephrine) in 95 patients within the referral population. The average number of 24-hour urine collections per patient was 1.36. Hereafter, only the first test after 1<sup>st</sup> January 2003 for each patient is considered. Of patients with high metanephrine or normetanephrine:

- 30 had no further follow-up because repeat testing of either urinary fractionated metanephrines or urinary free catecholamines was negative;
- 28 patients went on to imaging (at least an abdominal CT or MRI; in one case, an abdominal ultrasound only), from which 14 were diagnosed with pheochromocytoma, 9 with other pathologies (metastatic carcinoma, adrenal hemorrhage, adrenal myelolipoma, pelvic fibromatosis or neuromyotonia, which was diagnosed following neurophysiology tests) and 5 were deemed normal;
- 22 patients had been tested during an intercurrent condition (including intracranial hemorrhage, congestive cardiac failure, pancreatitis, sepsis, post-operatively from non-abdominal surgery or pregnancy); in 3 of these, subsequent repeat tests were

normal and there was no further follow-up; the remainder either died of their other pathology, or their symptoms resolved with no further follow-up;

- in 7 patients, no further follow-up of the abnormal result could be identified;
- in 5 patients the abnormal results were attributed to drug interference or effects (monoamine oxidase inhibitors, imipramine or poly-drug overdose).

Of the 14 patients in whom pheochromocytoma was detected, 11 were referred from hospital wards, 2 were referred from hospital clinics and only 1 was requested by a general practitioner.

A search of local histopathology records identified 22 pheochromocytomas of which 14 had had pre-surgery urinary metanephrines measured during the study period. In all of these 14 cases, the urinary metanephrines were abnormal. In 2 of these 14 cases, a subsequent diagnosis of multiple endocrine neoplasia was made. In addition, 4 paragangliomas were identified (a carotid body tumor, a left renal hilum paraganglioma in the context of Carney's triad, a urinary bladder paraganglioma and a jugulotympanic paraganglioma). In only one of these cases were urinary metanephrines measured pre-operatively: normetanephrine and metanephrine were normal, while 3-methoxytyramine was elevated.

Using our current laboratory thresholds, and reporting tests abnormal when either metanephrine or normetanephrine were high, the sensitivity for detecting pheochromocytoma was 100% with a specificity of 95%. However, because the pre-test probability in this population was very low, the positive predictive value was only 16%. In order to investigate better ways of interpreting the existing laboratory data, the distribution of metanephrines in the presence and absence of pheochromocytomas, for those patients in whom no previous pheochromocytoma or paraganglioma had been detected, was examined (Fig. 1). For most patients with pheochromocytoma, both metanephrine and normetanephrine were elevated. However, if the test of abnormality were to be "both metanephrine and normetanephrine high" then 2/14 cases of pheochromocytoma would have been missed (as one patient with a pheochromocytoma had a normal metanephrine, and another had a normal normetanephrine).

A simple measure that incorporates information about both metanephrine (MAD) and normetanephrine (NMT) is the product of their normalized (to correct for age and sex) metanephrine and normetanephrine values (nMAD.nNMT). In our population, the smallest nMAD.nNMT in which pheochromocytoma was histologically diagnosed was 3.1. Using a threshold of nMAD.nNMT > 3 as a test for pheochromocytoma yielded a test with sensitivity 100%, specificity 99.1% and positive predictive value 82%. To compare alternative measures for the detection of pheochromocytoma in our referral population, and to see whether the normalized 3-methoxytyramine concentration (n3MT) could add additional discriminative information when combined in a simple product, the receiver-operator characteristic curves (ROCs) were compared (Fig. 2). The area under the curves for the tests were: nMAD, 0.9970; nNMT, 0.9881; n3MT, 0.76; nMAD.nNMT, 0.9995; nMAD.nNMT.n3MT, 0.9968. It has been common practice to report a test positive if either nMAD or nNMT is abnormal. This can be assessed, using a modification of the technique used by Lenders et al. (11), by calculating  $k = \text{MAX}(nMAD, nNMT)$  (i.e. the maximum of nMAD and nNMT) and testing how the sensitivity and specificity varies as the threshold value  $k$  varies; this gave an area under the ROC curve of 0.9983. Finally, the sum of fractionated metanephrines (nMAD+nNMT) yielded an area under the ROC curve of 0.9989. The area under the ROC curve for nMAD.nNMT was the highest of all tests investigated and was statistically significantly different from each of the other six tests described above ( $P < 0.05$ ).

In all 24-hour urinary metanephrine tests, 3-methoxytyramine was also tested. 52 patients in the target population had an isolated elevation of 3-MT, of whom:

- 22 had normal repeats (or normal free urinary dopamine) and were not further investigated;
- 15 had no identifiable follow-up;
- 5 went on to further imaging; 3 studies were normal, two reported 'adrenal masses' which were both excised; one mass was a benign neurofibroma and the other was not adrenal, but a left renal hilum paraganglioma (mentioned previously);
- 4 were tested in the presence of known pathology (recurrent chemodectomas, or a previous pheochromocytoma, or MIBG-hot multiple liver lesions);
- 4 were tested during a severe intercurrent illness, of whom two retested normal;
- 2 were being treated with L-dopa.

## Discussion

Detecting rare problems is difficult, and is made even more difficult when the population tested is diverse and unselected. By examining the outcomes of those patients referred to our service for the testing of urinary metanephrines, we hope to aid clinical decision making by determining the causes commonly associated with abnormal test results and to optimize the use of available laboratory data. Of those patients with abnormal urinary metanephrines, about a third were considered to be false positives on the basis of repeat testing, a third went on to imaging studies, and in the other third the abnormal results were attributed to an acute intercurrent condition or the effects of drugs. It is interesting that even with abnormal urinary metanephrines and abnormal adrenal imaging, several other pathologies were found, including metastatic carcinoma, adrenal hemorrhage and adrenal myelolipoma. Hence, the combination of positive urinary metanephrines and positive adrenal imaging is not an appropriate gold standard for detecting pheochromocytomas.

The present work demonstrates that the normalized product of urinary metanephrine and normetanephrine (nMAD.nNMT) provides a useful measure for the detection of pheochromocytoma in the unselected referral base of a tertiary laboratory. There have been several other approaches used to increase the sensitivity or specificity of urinary metanephrines, including: careful analysis of the drug history (taking particular note of phenoxybenzamine, tricyclic antidepressants and beta blockers)(12); the absence of clonidine suppression of plasma normetanephrine (12-14); the use of over-night (12 hour) testing with a view to avoiding the variability introduced by day-time physical activity and to increase specificity because of the lower normal resting sympathetic activity over-night (15-17); special reporting for samples referred from ITU (18); or, the use of immunoassays for metanephrines to reduce drug interference (19). Each of these approaches has, and will, find favor in different situations depending on the pre-test probability and the mix of pathology within the referral population. The practical advantages of using nMAD.nNMT are that it requires no additional testing and its calculation can be readily automated.

Perhaps the strongest study supporting the use of plasma free metanephrines over urinary fractionate metanephrines is that of Lenders and colleagues (3). This predominantly prospective study was of a group of 1003 patients (over 7 years from 4 referral centers). Given that in this population, pheochromocytoma was found in 214 patients, the mean pre-test probability of greater than 20% makes this population very different from our hospital referral base. Indeed, the frequency of occurrence of pheochromocytoma in a screened hypertensive population is 0.5% (20), in those with adrenal incidentaloma it is 5% (21) and

42% in those presenting with MEN 2A and medullary thyroid cancer (22). In our unselected hospital referral population the mean pretest probability for pheochromocytoma was 0.8% (14/1819), or more than an order of magnitude lower than that described by Lenders and colleagues (3). In such low pretest probability populations, nMAD.nNMT provides excellent specificity while maintaining sensitivity, suggesting that this derivative measure could be used more widely as a screening test for pheochromocytoma.

### Study Limitations

The standard chosen for detecting pheochromocytomas was biopsy or excision-biopsy histopathology. While this method is very specific it is likely that small pheochromocytomas may have been missed and hence the sensitivity of the assays in the present work may be overestimated. This is particularly the case because in those patients with catecholamine-secreting tumors there is a strong correlation between urinary metanephrines (unfractionated) and tumor size (23). Even more significantly, the results of urinary fractionated metanephrines were used in this centre as part of the work-up for tumor excision; this biases our analysis towards higher sensitivity, particularly for the test 'nMAD or nNMT abnormal' (which is the test that generates an abnormal report from our laboratory). Despite this, the measure of nMAD.nNMT was better at predicting subsequent pheochromocytoma. The alternative gold standard of 'adrenal mass' has been demonstrated in this study, and others, to lack specificity and it seems impractical, if not unethical, to subject all 1819 patients to a CT or MRI. A longer period of follow-up would assist this and all such studies. The sample size in this study (1819 patients) is relatively small compared to some multicenter trials and the follow-up time of 28 months may not have been sufficient to allow the detection of all pheochromocytomas, which could also caused an over-estimate of sensitivity. A prospective study of the use of nMAD.nNMT in the diagnosis of pheochromocytoma is now required.

### Acknowledgments

We gratefully acknowledge the support of Dr S. Manek for allowing access to the histopathology database and the staff of the Department of Clinical Biochemistry who performed the analyses. The Wellcome Trust supports KLB.

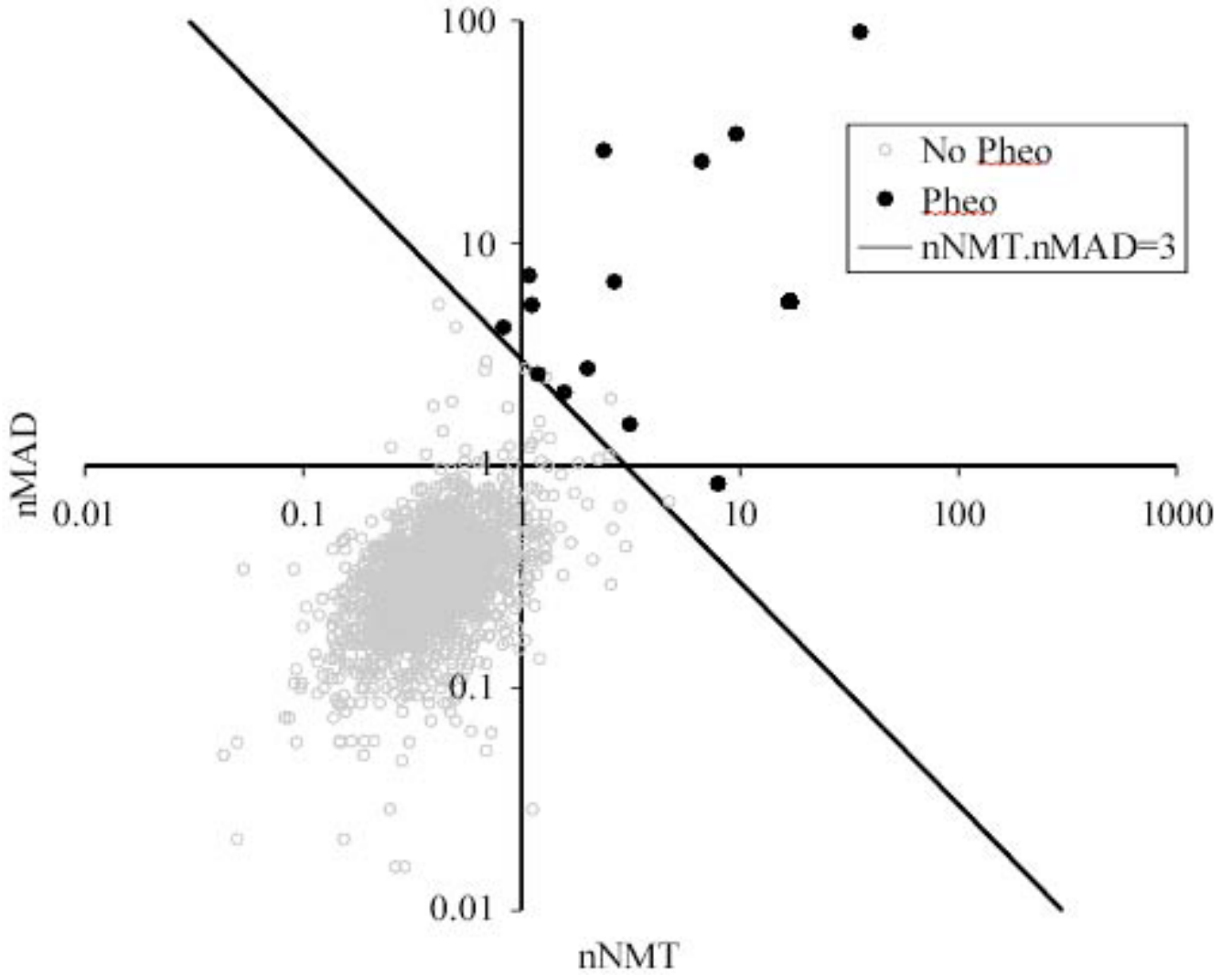
### Abbreviations

<b>NMT</b>	normetanephrine
<b>MAD</b>	metanephrine
<b>ROC</b>	Receiver-operator curve
<b>nNMT</b>	normalized normetanephrine
<b>nMAD</b>	normalized metanephrine
<b>n3MT</b>	normalized 3-methoxytyramine

### References

1. Stridsberg M, Husebye ES. Chromogranin A and chromogranin B are sensitive circulating markers for pheochromocytoma. *Eur J Endocrinol.* 1997; 136:67–73. [PubMed: 9037129]
2. Ueda E, Kuchii M, Hano T, Ueno Y, Jo Y, Shima H, et al. Platelet noradrenaline content as an integrated measure of variations in plasma noradrenaline. *Jpn Circ J.* 1990; 54:383–90. [PubMed: 2398618]
3. Lenders JW, Pacak K, Walther MM, Linehan WM, Mannelli M, Friberg P, et al. Biochemical diagnosis of pheochromocytoma: which test is best? *Jama.* 2002; 287:1427–34. [PubMed: 11903030]

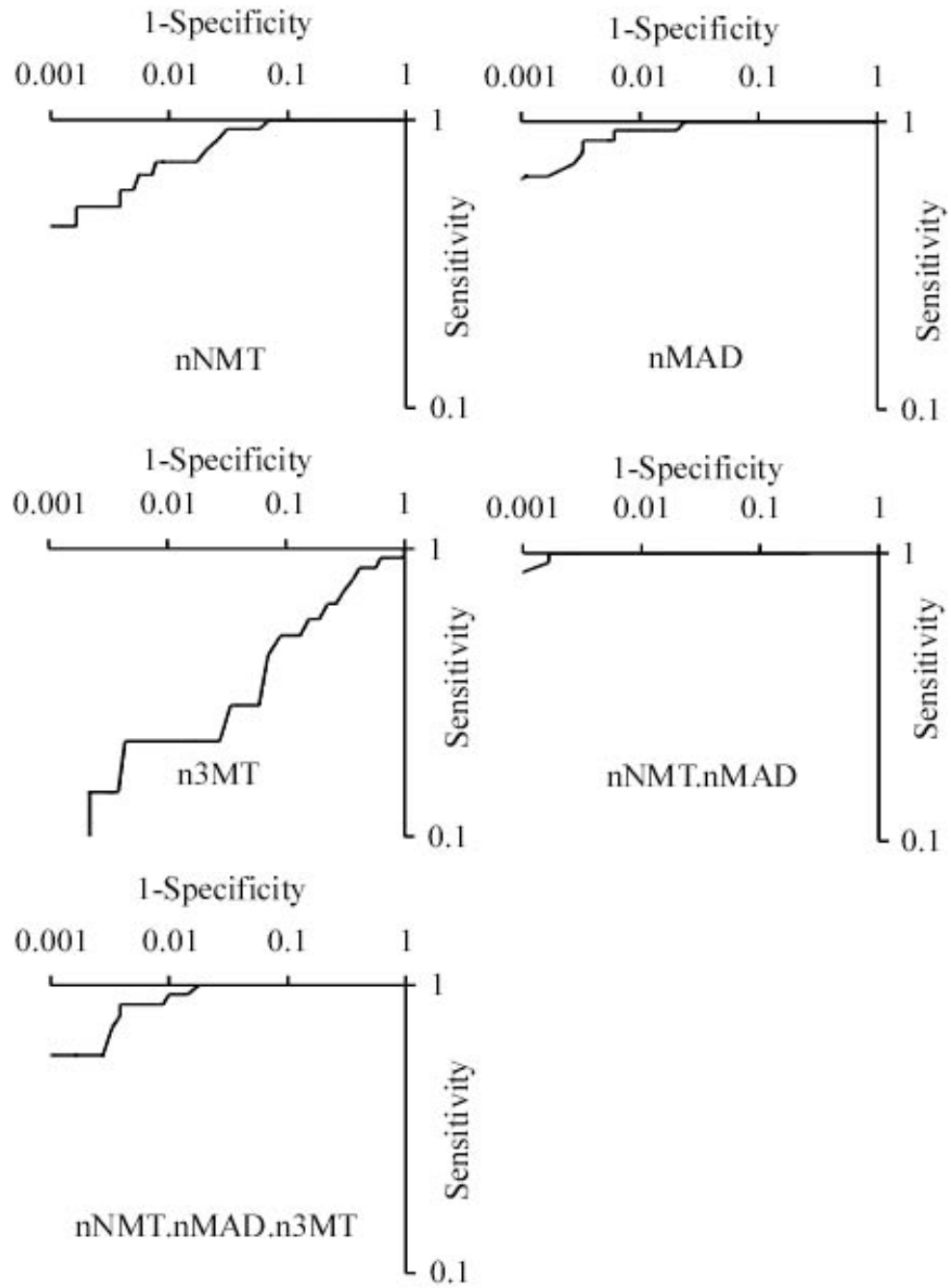
4. Pacak K. Pheochromocytoma. Accessed March 2006 <http://www.endotext.org/adrenal/adrenal34/adrenalframe34.htm>
5. Pacak K, Linehan WM, Eisenhofer G, Walther MM, Goldstein DS. Recent advances in genetics, diagnosis, localization, and treatment of pheochromocytoma. *Ann Intern Med.* 2001; 134:315–29. [PubMed: 11182843]
6. 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.* 2003; 88:553–8. [PubMed: 12574179]
7. Gardet V, Gatta B, Simonnet G, Tabarin A, Chene G, Ducassou D, Corcuff JB. Lessons from an unpleasant surprise: a biochemical strategy for the diagnosis of pheochromocytoma. *J Hypertens.* 2001; 19:1029–35. [PubMed: 11403350]
8. Eisenhofer G. Editorial: biochemical diagnosis of pheochromocytoma--is it time to switch to plasma-free metanephrines? *J Clin Endocrinol Metab.* 2003; 88:550–2. [PubMed: 12574178]
9. Standing SJ, Gardner SG, Shine B, Wass JAH, Turner HE. Reference ranges for urinary metadrenalines in patient populations. *Endocrine Abstracts.* 2002; 3:147.
10. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* 1983; 148:839–43. [PubMed: 6878708]
11. Lenders JW, Keiser HR, Goldstein DS, Willemsen JJ, Friberg P, Jacobs MC, et al. Plasma metanephrines in the diagnosis of pheochromocytoma. *Ann Intern Med.* 1995; 123:101–9. [PubMed: 7778821]
12. 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.* 2003; 88:2656–66. [PubMed: 12788870]
13. Grossman E, Goldstein DS, Hoffman A, Keiser HR. Glucagon and clonidine testing in the diagnosis of pheochromocytoma. *Hypertension.* 1991; 17:733–41. [PubMed: 2045133]
14. Taylor HC, Mayes D, Anton AH. Clonidine suppression test for pheochromocytoma: examples of misleading results. *J Clin Endocrinol Metab.* 1986; 63:238–42. [PubMed: 3711261]
15. Ganguly A, Henry DP, Yune HY, Pratt JH, Grim CE, Donohue JP, Weinberger MH. Diagnosis and localization of pheochromocytoma. Detection by measurement of urinary norepinephrine excretion during sleep, plasma norepinephrine concentration and computerized axial tomography (CT-scan). *Am J Med.* 1979; 67:21–6. [PubMed: 463912]
16. Macdougall IC, Isles CG, Stewart H, Inglis GC, Finlayson J, Thomson I, et al. Overnight clonidine suppression test in the diagnosis and exclusion of pheochromocytoma. *Am J Med.* 1988; 84:993–1000. [PubMed: 3376984]
17. Peaston RT, Lennard TW, Lai LC. Overnight excretion of urinary catecholamines and metabolites in the detection of pheochromocytoma. *J Clin Endocrinol Metab.* 1996; 81:1378–84. [PubMed: 8636337]
18. Smythe GA, Drew CM. REPCAT: desktop expert system for interpreting and validating laboratory data for pheochromocytoma diagnosis with the database application Omnis 7. *Clin Chem.* 1997; 43:134–40. [PubMed: 8990235]
19. Wassell J, Reed P, Kane J, Weinkove C. Freedom from drug interference in new immunoassays for urinary catecholamines and metanephrines. *Clin Chem.* 1999; 45:2216–23. [PubMed: 10585355]
20. Manger, WM.; Gifford, R, Jr.. *Clinical and experimental pheochromocytoma.* Cambridge, Mass.: Blackwell Science; 1996.
21. Young WF Jr. Management approaches to adrenal incidentalomas. A view from Rochester, Minnesota. *Endocrinol Metab Clin North Am.* 2000; 29:159–85. [PubMed: 10732270]
22. Howe JR, Norton JA, Wells SA Jr. Prevalence of pheochromocytoma and hyperparathyroidism in multiple endocrine neoplasia type 2A: results of long-term follow-up. *Surgery.* 1993; 114:1070–7. [PubMed: 7903003]
23. Eisenhofer G, Lenders JW, Goldstein DS, Mannelli M, Csako G, Walther MM, et al. Pheochromocytoma catecholamine phenotypes and prediction of tumor size and location by use of plasma free metanephrines. *Clin Chem.* 2005; 51:735–44. [PubMed: 15718487]



**Figure 1. The normalized referral population distribution of normetanephrine and metanephrine**

Each point represents the values of normetanephrine (NMT) and metanephrine (MAD), divided by the age- and sex- dependent upper limits of normal (9), in the first test after 1<sup>st</sup> January 2003 for each patient. Those with recurrent pheochromocytoma, or previously known von Hippel Lindau syndrome or neurofibromatosis were excluded. Those patients subsequently histologically diagnosed with pheochromocytoma are marked. A line is drawn to show a suitable detection threshold where  $nNMT \cdot nMAD = 3$ .





**Figure 2. Logarithmic ROC curves for 5 measures used to detect pheochromocytomas** Receiver-operator (ROC) curves for normalized urinary normetanephrine (nNMT), metanephrine (nMAD) and 3-methoxytyramine (n3MT), as well as the derivative measures nNMT.nMAD and nNMT.nMAD.n3MT, are shown. The gold standard used for diagnosis was histologically-confirmed pheochromocytoma before the 4<sup>th</sup> May 2005.