



Article

# Measurements of Plasma-Free Metanephrines by Immunoassay Versus Urinary Metanephrines and Catecholamines by Liquid Chromatography with Amperometric Detection for the Diagnosis of Pheochromocytoma/Paraganglioma

Wolfgang Raber <sup>1,\*</sup>, Hans Kotal <sup>1</sup>, Rodrig Marculescu <sup>2</sup> , Christian Scheuba <sup>3</sup>,  
Martin B. Niederle <sup>4</sup> , Alexandra Kautzky-Willer <sup>1</sup> and Michael Krebs <sup>1</sup>

<sup>1</sup> Division of Endocrinology & Metabolism, Department of Medicine III, Medical University of Vienna, 1090 Vienna, Austria; h.kotal@gmx.at (H.K.); alexandra.kautzky-willer@meduniwien.ac.at (A.K.-W.); michael.krebs@meduniwien.ac.at (M.K.)

<sup>2</sup> Department of Laboratory Medicine, Medical University of Vienna, 1090 Vienna, Austria; rodrig.marculescu@meduniwien.ac.at

<sup>3</sup> Department of Surgery, Medical University of Vienna, 1090 Vienna, Austria; christian.scheuba@meduniwien.ac.at

<sup>4</sup> Department of Anesthesia, General Intensive Care and Pain Therapy, Medical University of Vienna, 1090 Vienna, Austria; martin.niederle@meduniwien.ac.at

\* Correspondence: wolfgang.raber@meduniwien.ac.at; Tel.: +43-1-40400-43120

Received: 20 August 2020; Accepted: 24 September 2020; Published: 26 September 2020



**Abstract:** Studies conflict concerning the use of enzyme immunoassays (EIA) for plasma free metanephrines (P-MNs) vs. other methods for pheochromocytoma/paraganglioma (PPGL) diagnosis. We compared commercially available EIAs for P-MNs with high-pressure liquid chromatography (HPLC) for 24 urinary MNs (U-MNs) and -catecholamines (U-CATs). 943 (565 female, 378 male) patients (54 PPGL, 889 Non-PPGL) were studied. Simultaneous measurements of all parameters analyzed at the central lab of our university hospital was mandatory for inclusion. Sensitivity of P-MNs (94.4%) was similar to that of U-MNs (100%), and both were higher than of U-CATs (77.8%), specificity of P-MNs (100%) higher than of U-MNs (73.6%), and similar to U-CATs (99.8%). With the recently proposed downward adjusted ULN of P-MNs to correct for the reported negative bias of the EIAs sensitivity (98.1%) raised non-significantly, but specificity decreased significantly (94.8%). Areas under receiver-operating characteristic curves indicated comparable diagnostic performance of P-MNs (0.989) vs. U-MNs (0.995), both better than U-CATs (0.956). In summary, the EIAs to measure P-MNs performed similarly to U-MNs by HPLC, and both better than U-CATs by HPLC. The post-test probability of PPGL given a positive test result was best for P-MNs, and higher than for the other pairs of analytes. Downward corrections of ULN of P-MNs did not improve test performances.

**Keywords:** pheochromocytoma; diagnosis; plasma metanephrines; immunoassay; urinary metanephrines; urinary catecholamines

## 1. Introduction

Pheochromocytomas (PHEO) and paragangliomas (PGL; together: PPGL) are rare catecholamine-producing tumors estimated to occur in about 2–8 out of 1 million people per year [1]. The prevalence in the hypertensive population is 0.1–0.5% [1–3], in an unselected hospital referral population 0.8–1.6% [4–6], and in patients with adrenal incidentalomas about 5% [7]. Those with a history of PPGL have a

risk of recurrence as high as 16.5% [8], and 30–40% of PPGL may be associated with a hereditary disease [1,9], all scenarios typical of a large referral hospital and a specialized endocrine unit, respectively. The diagnosis may be challenging as about 50% of patients are asymptomatic [10] or suffer from only unspecific signs or symptoms [9,11]. On the other hand, up to 20% may develop life-threatening hemodynamic instability leading to pulmonary edema, malignant arrhythmias, myocardial infarction, stroke, or sudden death [12,13].

The 2014 Endocrine Society Clinical Practice Guideline with a level A quality of evidence by the GRADE system [14] recommends that initial biochemical testing should include the measurement of plasma-free or urinary fractionated metanephrines (MNs) with no preference of one over the other. The panel gives a grade 2 (weak) recommendation to use liquid chromatography (LC) with mass spectrometric (MS) or electrochemical detection (ED) rather than other lab methods [15]. Nonetheless, even in university hospitals, commercially available enzyme-linked immunoassays (EIAs) for the determination of plasma metanephrine (MN) and normetanephrine (NMN) are in widespread use due to their low price, ease of handling and—compared to chromatographic/MS methods—greater likelihood of being certified by ISO (International Organization for Standardization, Geneva, Switzerland) norms.

To date, no laboratory of the health care system of the Republic of Austria, in 2020 serving an estimated population of 8.9 millions (<https://en.wikipedia.org/wiki/Austria>), has been offering chromatographic/MS methods for quantification of plasma-free metanephrines for routine patient care, neither in private institutions nor in university hospitals. Capital costs of instrumentation, the level of operator expertise, and the need for in-house validation are higher than with EIA methods [16]. It is of note that in the European Union, until only recently, there has been no single officially CE-IVD (Conformité Européenne, In-Vitro Diagnostic Medical Device)-approved mass spectrometric or coulometric detector for the high-pressure (HP)LC to quantify plasma-free metanephrines—however, this improvement is mandatory for many labs due to regulatory legal requirements for the implementation of new methods. Based on the non-availability of these methods, we report our experience of  $n = 943$  patients tested for PPGL in routine clinical care, in every patient and simultaneously using the three test methods routinely available (plasma-free metanephrines by EIA, 24-h urinary fractionated metanephrines, and 24-h urinary catecholamines, the latter two by HPLC with amperometric detection), compute which performs best for which metabolites to be analyzed, and compare the methods between each other. We think that the question of whether state-of-the-art care is possible without HPLC-MS/MS is of interest for analysts in all countries, where HPLC-MS/MS is not available. In 2015, the commercially available EIAs of Labor Diagnostika Nord, Nordhorn, Germany (LDN) have been reported to systematically underestimate plasma-free metanephrines contributing to a poor diagnostic sensitivity of only 74.1% in one series [17]. We thus wanted to know whether the correction algorithm, as proposed by Weismann et al. [17] to overcome their published negative bias of the EIAs suggested to have caused the low diagnostic sensitivity [17], is of clinical importance in a different population, or whether patient care would be possible without using these downward corrected upper limits of normal (ULN) of plasma-free metanephrines.

The study aims were: (1) Calculation of the diagnostic accuracy of the metabolites outlined in Table 1 as measured by the diagnostic sensitivity and specificity, by receiver–operating curve (ROC)-based analysis of the area under the curve (AUC) of single parameters, and of the pairs of measurements, and by using the posterior probability of PPGL given a positive and negative test result of the pairs of measurements, when upper limits of normal (ULN) stipulated by the manufacturers are applied; (2) assessment of the diagnostic accuracy of the EIAs following use of ULN to correct for the described negative bias as proposed recently [17], and comparison with the determination of the urinary metabolites.

**Table 1.** Upper limits of normal (ULN) of the metabolites as used in the analyses of this study.

| Biochemical Parameter | ULN         |
|-----------------------|-------------|
| P-MN                  | 90 pg/mL    |
| P-NMN                 | 180 pg/mL   |
| P-MN *                | 58 pg/mL    |
| P-NMN *               | 45–88 pg/mL |
| U-MN                  | 138 µg/24 h |
| U-NMN                 | 311 µg/24 h |
| U-A                   | 27 µg/24 h  |
| U-NA                  | 97 µg/24 h  |

The asterisks (\*) denote P-MN and P-NMN with the corrected ULN as proposed by others [17]: In short, the two values for P-NMN \* are the lowest and highest concentrations of the age-adjusted ULN, from 45 pg/mL at the age of 5 years to a maximum at the age of 65 years and above. Details are given in the legend of Table 3 of their work [17]. To convert values for plasma measurements to picomoles per liter, and for urinary measurements to nmol/day, multiply by 5.46 for normetanephrine, 5.08 for metanephrine, 5.91 for norepinephrine, and 5.46 for epinephrine. Please note that after the study has been completed, the lab has issued new ULN for P-MN and P-NMN (68 and 196 pg/mL, respectively), and for U-MN and U-NMN (311 and 390 µg/24 h, respectively), but patient care and all analyses were by the above ULN, respectively.

## 2. Subjects and Methods

### 2.1. Subjects

The database consisted of 3373 patients from the Department of Laboratory Medicine of the Vienna General Hospital (in short, the lab). Clinical, biochemical, radiological, and histological data were obtained retrospectively. A mean of 1.4 (range: 1–5) samples for each plasma and 24 h-urinary measurements were obtained in patients with PPGL before surgery. Hereafter, only the first test results of the patients are considered for analysis.

The Vienna General Hospital generally serves as a tertiary referral center. Still, given the lack of insurance constraints in our country, many patients choose to attend the specialized university out-patient services for initial check-up of their complaints. Thus, for the purpose of this study, a low pre-test probability of PPGL can be reasoned. Two-thirds of the 3373 patients were seen either at the specialized endocrine out-patient unit or treated for various conditions as an in-patient. No attempt was made to differentiate between out-patients and those admitted to the hospital, nor to evaluate the medications taken by the patients at the time of biochemical evaluation. Only patients with all plasma and 24 h-urinary measurements, determined in-house and at the same time (with a tolerance of a maximum of one week apart), were included (Figure 1).

The indications for biochemical testing were assessed in 683 patients presenting with at least one positive measurement of any parameter: Hypertension resistant to  $\geq 3$  antihypertensive drugs or severe hypertension, or of young age ( $n = 346$ ), the finding of an incidental adrenal mass ( $n = 120$ ), symptoms suggestive of PPGL such as sweating, palpitations, or headache with or without paroxysmal hypertension ( $n = 101$ ), suspected or confirmed genetic syndromes [1,9] associated with PPGL ( $n = 100$ ), and a previous history of PPGL ( $n = 16$ ). No effort was made to evaluate the indications for testing in the 260 patients presenting with negative measurements of all analytes at all times during follow-up (FU). Exclusion criteria are given in Figure 1.

$N = 943$  patients were thus included for analyses (Figure 1 and Table 2). Biochemical assessment was performed a mean of 7.6 times (range: 2–15) in patients assigned PPGL negative during the FU period of at least 2 years. Negative biochemical tests of all parameters at all times (from first presentation until  $\geq 2$  years of FU) were identified in  $n = 260$  patients and were categorized as PPGL negative (PPGL-). Positive findings on at least one occasion of at least one analyte, (i) above the ULN stipulated by the manufacturer of the EIAs and the HPLC, respectively, or (ii) above the downward corrected ULN of P-MNs as proposed recently [17], were identified in  $n = 683$  patients. The diagnosis of PPGL (PPGL+) was defined by histological criteria ( $n = 52$ ), or by positive  $^{123}\text{I}$ -MIBG scan ( $n = 1$ ) and positive F-DOPA PET-CT ( $n = 1$ ) of two older patients presenting with grossly elevated concentrations

of P-MNs, U-MNs, and U-CATs but refused surgery.  $N = 629$  patients with at least one positive biochemical result were categorized PPGL- by best possible criteria (Table 2): By surgical exclusion of PPGL in  $n = 2$  patients (one adrenal Cushing-syndrome and one adrenal adenoma), by normal thoraco-abdominal imaging (CT or MRI) and normal laboratory FU for at least 2 years ( $n = 158$ ), by consistently normal biochemical findings together with absence or remission during a FU of  $\geq 2$  years of symptoms or signs suggestive of PPGL ( $n = 190$ ), by an alternative diagnosis such as panic disorder or withdrawal from multiple substance abuse ( $n = 10$ ), and by an uneventful clinical FU of  $\geq 2$  years including lack or remission of symptoms or signs suggestive of PPGL, and lack of perioperative complications during surgery for other conditions ( $n = 269$ ).

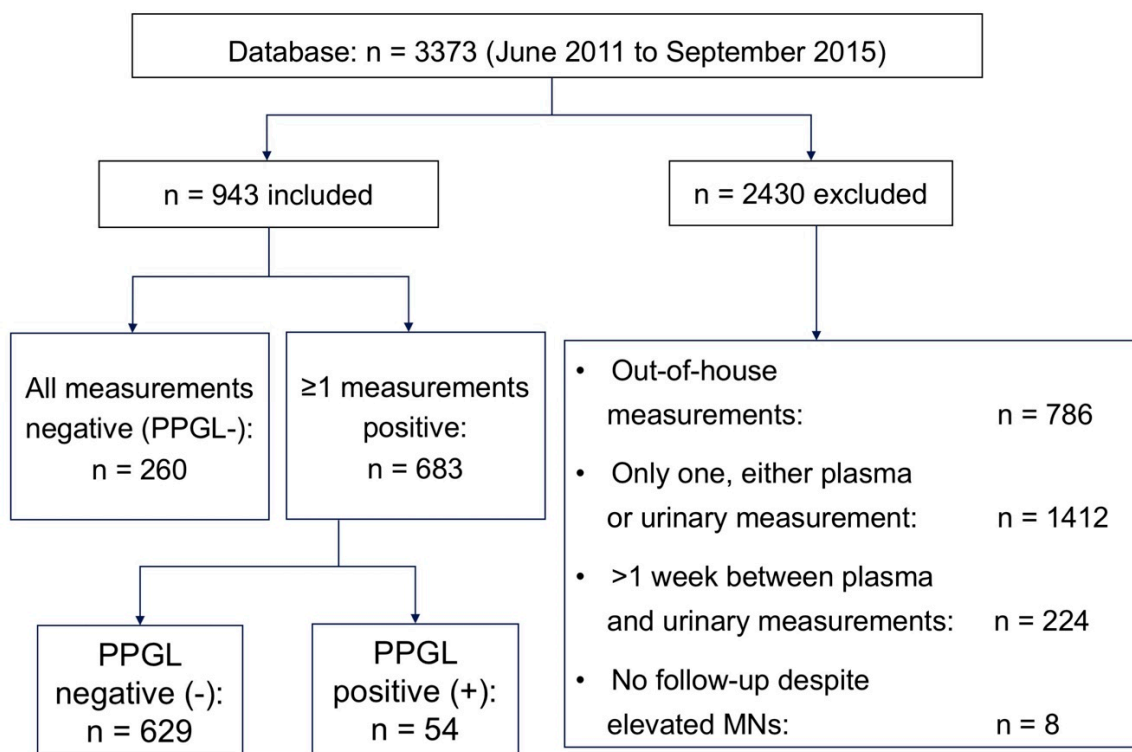


Figure 1. Flow-chart of patients.

Table 2. Patients with at least one parameter above the lowest respective upper limits of normal (ULN) (see Table 2) and allocated to correct positive (PPGL+) and correct negative (PPGL-) by best possible criteria as outlined in the text.

| <i>n</i> | PPGL+   | N   | PPGL-                 |
|----------|---|-----|-----------------------|
| 52       | histology+  | 2   | histology-            |
| 2        | lab+ and J- MIBG-scintigraphy+ (1) and F-DOPA PET-CT+ (1) | 158 | lab and radiol FU-    |
|          |   | 190 | lab and clinical FU-  |
|          |   | 10  | alternative diagnosis |
|          |   | 269 | clinical FU-          |

Please note, that the  $n = 260$  patients with all parameters below the respective ULN (for P-MN and P-NMN: Both had to be below the respective downward corrected ULN [17]), add to the total number of PPGL- patients.

## 2.2. Methods

All analyses were performed in an ISO 9001 certified and ISO 15189 accredited laboratory at the Department of Laboratory Medicine, Medical University of Vienna. All methods were approved for clinical diagnostic use according to the European Union In Vitro Diagnostics Directive 98/97/EG Conformité Européenne, In-Vitro Diagnostic Medical Device (CE IVD) labeled and were performed

according to the manufacturers' instructions for use. Analytical performance was constantly monitored by participation in international proficiency testing programs from INSTAND (Düsseldorf, Germany) and The Royal College of Pathologists of Australasia (RCPI, Surry Hills, NSW, Australia). In addition, the laboratory regularly exchanged samples and compared results with the Central Institute for Medical and Chemical Laboratory Diagnostics, Medical University of Innsbruck, Innsbruck, Austria according to Clinical and Laboratory Standards Institute [CLSI] Guideline QMS24. Plasma MNs were measured by the commercially available 2-MET Plasma ELISA Fast Track Assay kit of LDN [18] including two separate assays for P-MN and P-NMN, respectively. Since February 2013, all blood samples have been drawn according to standard operating procedures requiring a prior overnight fast and a supine resting position for at least 30 min. Before that date, about 70–80% of blood samples—corresponding to the percentage of patients from the endocrinology department—were obtained that way. No information was available of the other senders before February 2013. All blood samples were collected into K3EDTA tubes, plasma was separated by centrifugation, and specimens were stored at  $-20^{\circ}\text{C}$  before analysis. Details are described elsewhere [16,17]. The intra-assay coefficients of variation (CV) ranged from 7.3 to 22 % for P-MN, and 7.2 to 7.8 % for P-NMN, the inter-assay CV from 12.0 to 21.1% for P-MN, and 11.7 to 13.0% for P-NMN. The lower limits of quantification (LLOQ) were 17 pg/mL (0.086 nmol/l) for MN, and 23 pg/mL (0.126 nmol/l) for NMN, respectively. The ULN were derived from fasting blood samples of 96 healthy subjects of the red cross Ljütensee near Hamburg, Germany (GER), and obtained in the sitting position (personal communication, Klemm R, Sales & Marketing Manager, LDN). The 24-h urinary measurements were by HPLC with amperometric detection using commercial kits of Chromsystems, Munich, Germany, as described previously [19]. The inter-assay CV ranged from 8.0 to 11.0% for NMN and MN, and from 5.0 to 8.0% for the CATs. Within-assay CV were <10% for both compounds. The LLOQ were 33, 5, 6, and 1  $\mu\text{g}/24\text{ h}$  for NMN, MN, NA, and A, respectively.

Concentrations of test parameters are given as percent (%) of ULN to allow for easy comparisons between the methods. The ULN used for analyses are given in the Table 1.

### 2.3. Statistics

Sensitivity (defined as the fraction of patients with the disease identified as positive by the test), and specificity (fraction of those without the disease correctly identified as negative by the test) of the biochemical parameters were calculated for each analyte alone and as pairs (combined model). For the combined model, a positive finding was defined as a concentration above the ULN of at least one analyte, MN or NMN and A or NA, respectively. For a negative result, the concentrations of both analytes had to be below the respective ULN. Ninety-five percent confidence intervals (CI) of sensitivities and specificities were computed by the method of Wilson [20] and compared by McNemar's test using a chi-square approximation [21].

To evaluate the diagnostic performance independent of ULN, ROC curves based on logistic regression analysis were created for each analyte as well as for the pairs, and 95% CI calculated by the method of Newcombe [22]. The ROC curves as a summary measure of the clinical utility of the tests show the relative changes in rates of true-positive (sensitivity plotted on the y-axis) and false-positive results (1-specificity plotted on the x-axis) for the diagnosis of PPGL as a function of the respective ULN. To test the null hypothesis of equal diagnostic performance of EIA and HPLC, the method of Hanley and McNeil [23] was applied. When the critical value Z of the standard normal distribution exceeded  $\pm 1.96$  of the two-sided significance of 0.05, the null hypothesis of equal performance was rejected.

To determine the probability of a patient suffering from PPGL given a positive and a negative test result, Fagan's nomogram for Bayes theorem based on the pre-test probability (numbers of PPGL divided by the total study population) and the diagnostic sensitivity and specificity was used [24,25]. Post-test probabilities (post-TP) with 95% CI were assessed with the diagnostic test calculator software version 201004210, free software available under the Clarified Artistic License [26]. All other computations were performed using GraphPad Prism version 8.0.0 for Mac, GraphPad Software, San Diego, CA, USA, [www.graphpad.com](http://www.graphpad.com). Means and medians were compared by the t-test

and the Mann-Whitney U test (both two-sided), respectively, as appropriate. For the comparison of three or more means, one-way ANOVA and Tukey's multiple comparisons test was used, for non-parametric comparisons of three or more medians the Kruskal–Wallis and Dunn's multiple comparisons test. *P*-values < 0.05 were considered significant.

The study has been approved by the ethics committee of the Medical University of Vienna (No. 1089/2016).

### 3. Results

#### 3.1. Baseline Data

Age and sex were comparable between PPGL+ and PPGL− patients (*p* = n.s.). A genetic disease was identified in 14 of the 54 patients with PPGL (seven with multiple endocrine neoplasia type 2, two with von Hippel-Lindau disease, two with neurofibromatosis, two with succinate dehydrogenase mutation B, and one with PGL-3 locus mutation). These patients were younger (*p* = 0.001) than those with sporadic disease (37 ± 12 vs. 53 ± 14 years), but concentrations of biochemical parameters were comparable (*p* = n.s.) to patients with sporadic tumors. There were 5 patients with PGL (2 of them with metastases), 2 patients with both PGL and PHEO, and 49 PHEO among the 54 patients with confirmed PPGL (Table 3). As to the possible distinctive biochemical profile of patients with adrenal (PHEO) versus extra-adrenal (PGL) tumor, parameters were analyzed separately for both entities. The latter displayed higher (*p* = 0.04 and 0.03, respectively) U-NMN and U-NA, but similar concentrations of the other parameters (Supplementary Figure S1). The small number (*n* = 5) of PGL precludes further evaluation as to the accuracy of the diagnostic tests in PGL vs. PHEO.

**Table 3.** Characteristics of the study population of *n* = 943 patients and concentrations of biochemical parameters (median and absolute range).

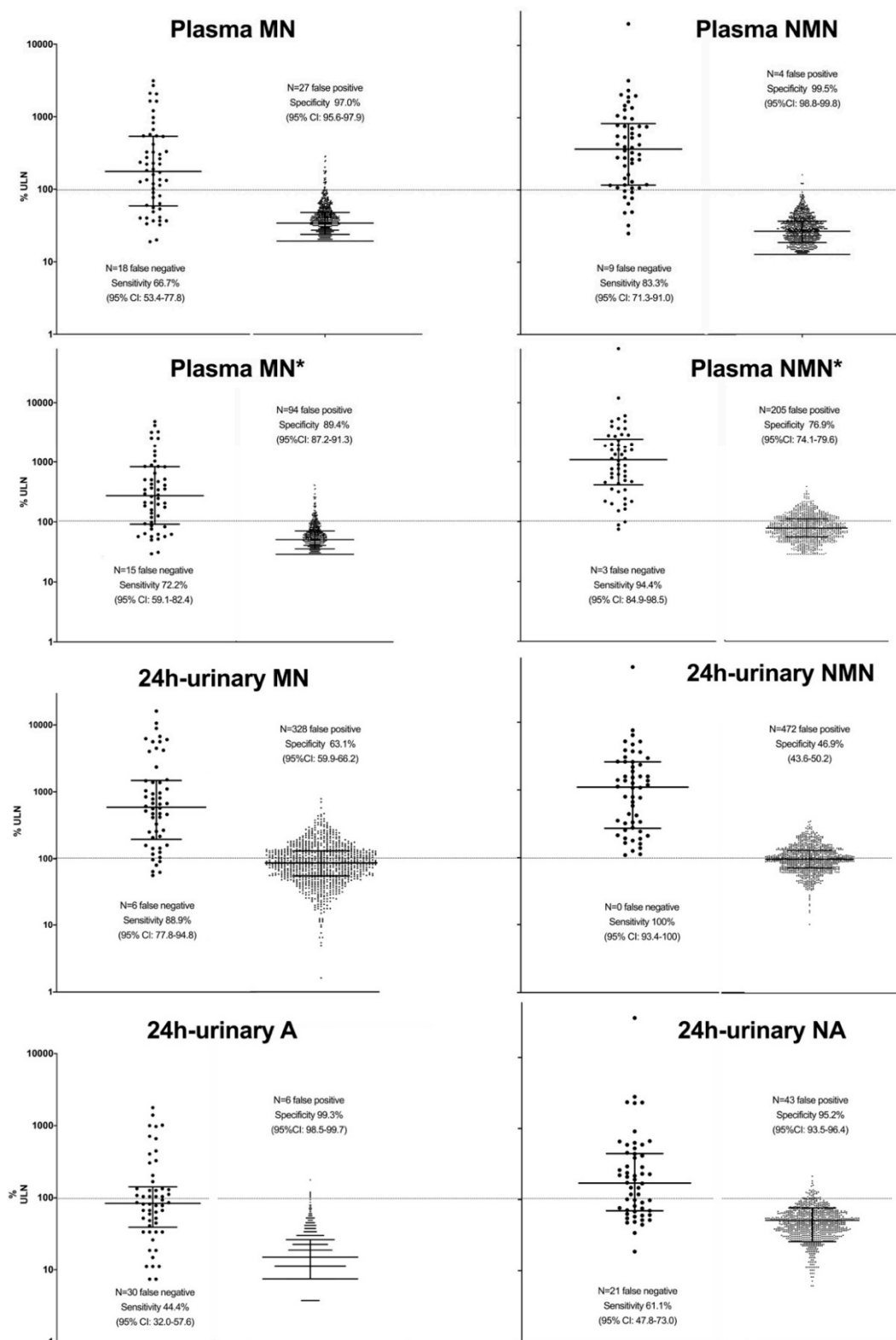
| Patients                       | <i>n</i> = 54 PPGL+ | <i>n</i> = 889 PPGL− | <i>p</i> -Value |
|--------------------------------|---------------------|----------------------|-----------------|
| Age (years)                    | 51 (18–76)          | 55 (5–90)            | 0.55            |
| Female sex, <i>n</i> (%)       | 28 (52)             | 537 (60)             | 0.80            |
| Genetic syndrome, <i>n</i> (%) | 14 (26)             | 0                    |                 |
| P-MN (pg/mL)                   | 159 (17–2815)       | 30 (17–248)          | 0.001           |
| P-NMN (pg/mL)                  | 663 (45–35,832)     | 48 (23–290)          | 0.001           |
| U-MN (μg/24 h)                 | 814 (78–22,248)     | 118 (5–672)          | 0.001           |
| U-NMN (μg/24 h)                | 3426 (340–203,009)  | 321 (33–1388)        | 0.001           |
| U-A (μg/24 h)                  | 23 (2–480)          | 4 (1–47)             | 0.001           |
| U-NA (μg/24 h)                 | 166 (18–34,987)     | 46 (6–314)           | 0.001           |

Please note the LLOQ of 17 and 23 pg/mL for plasma-NMN and -MN, and of 33, 5, 6, and 1 μg/24 h for 24 h-urinary NMN, -MN, -NA, and -A, respectively. To convert values for plasma measurements to picomoles per liter, and for urinary measurements to nmol/day, multiply by 5.46 for normetanephrine, 5.08 for metanephrine, 5.91 for noradrenaline, and 5.46 for adrenaline, respectively.

#### 3.2. Sensitivity and Specificity

##### 3.2.1. Sensitivity and Specificity of Single Parameters

U-NMN detected all 54 PPGL and provided higher (*p* < 0.05) sensitivity than all other single parameters, except for P-NMN (51 of 54 PPGL positive) with the corrected ULN as proposed by others [17], U-A and U-NA the lowest. On the other hand, P-NMN was most specific (4 of 889 individuals without PPGL were false positive), more (*p* < 0.05) than all other single metabolites except for U-A (six false positives). U-NMN (472 false positives) performed worst regarding specificity. Figure 2 shows the sensitivity and specificity of the analytes according to the ULN given in Table 1. Comparisons of sensitivities and specificities including 95% CI are given in Supplementary Table S1.



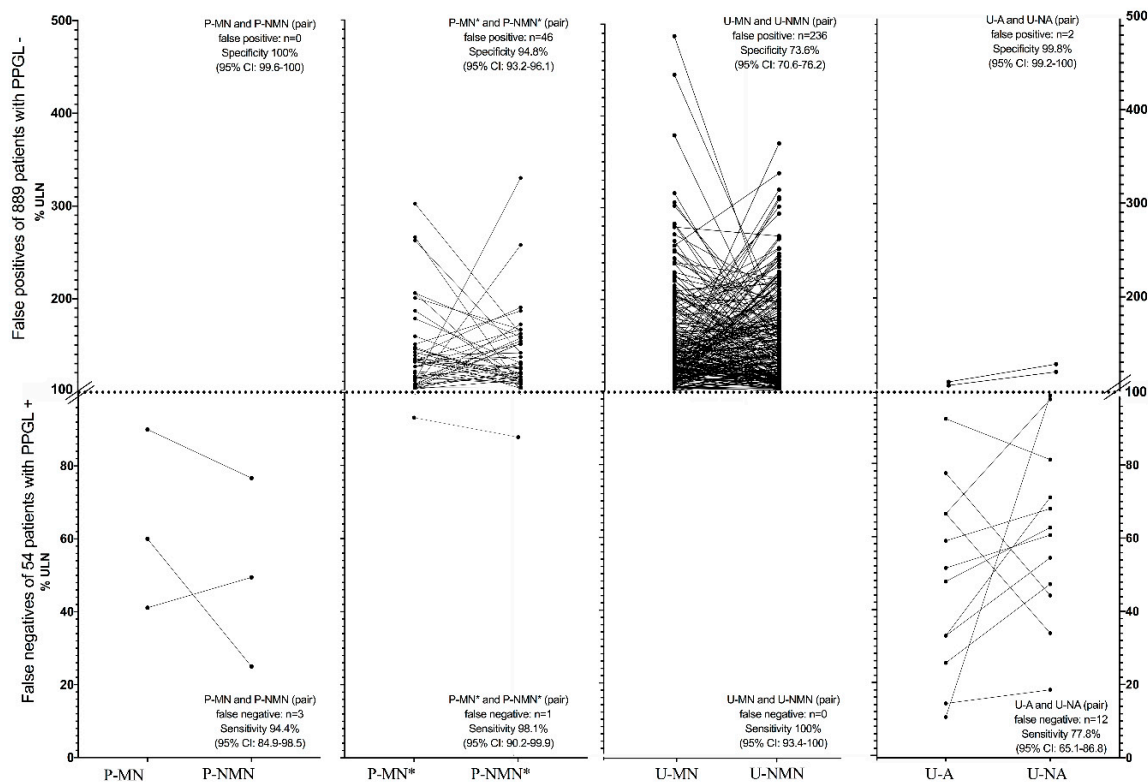
**Figure 2.** Sensitivities and specificities of single parameters with all ULN. Data are presented for 54 patients (left columns), and for 889 individuals without (right columns) PPGL, respectively. Concentrations are expressed as percentages of the respective ULN (dashed line). Lines within the graph depict medians and inter-quartile ranges. The asterisks (\*) denote P-MN and P-NMN with the corrected ULN as proposed recently [17]. Please note the logarithmic scale!

### 3.2.2. Sensitivity and Specificity of Parameters Combined

Clinical decisions are seldom based on a single parameter, so sensitivities and specificities were computed of combined models (pairs) as well. The sensitivity (Figure 3, below the dashed line) of the pair of P-MN and P-NMN (51 of 54 PPGL detected) was higher ( $p \leq 0.02$ ) than that of the pair of U-A and U-NA (42 of 54 PPGL detected), but not ( $p = n.s.$ ) vs. the combined U-MN and U-NMN (all 54 PPGL detected). Specificity (Figure 3, above the dashed line) was highest with the combined P-MN and P-NMN (no false positives of 889 patients without PPGL), and higher ( $p < 0.001$ ) than that of the pair of 24 h-urinary metanephrines (236 false positive) and of plasma-free metanephrines (46 false positive), when using the corrected ULN as proposed by others [17].

As depicted in Figure 3, using this downward corrected ULN of plasma-free metanephrines as suggested by Weismann et al. [17] resulted in a statistically non-significant increased sensitivity versus no corrections of ULN (one vs. three false negative), at the cost of a significant decrease in specificity (0 vs. 46 false positive).

False positive results of U-MN and U-NMN showed a broader range ( $p < 0.05$ ) up to almost 5-fold the ULN when compared to the range of false positive concentrations (up to 3.3-fold the ULN) of P-MN and P-NMN with the corrected ULN [17]. Comparisons of sensitivities and specificities including 95% CI are summarized in Table 4.



**Figure 3.** False negative results of 54 patients with PPGL+ (below dotted line depicting the ULN) and false positive results of 889 patients with PPGL- (above dotted line) of the combined parameters (P-MN+P-NMN, U-MN+UNMN, U-A+U-NA). Intraindividual measurements within the pairs of parameters are connected with solid lines. The asterisks (\*) denote P-MN and P-NMN with the corrected ULN as proposed by others [17]. Please note the different scale below and above the dotted line.



**Table 4.** Sensitivity (above) and specificity (below) of the pair of plasma compared to the pair of 24 h-urinary metabolites, respectively.

| Comparison EIA vs. HPLC              | EIA (Plasma) Sensitivity (95% CI); ULN | HPLC (24 h-Urine) Sensitivity (95% CI); ULN | p-Value      |
|--------------------------------------|--|---|--------------|
| P-MN + P-NMN/U-MN + U-NMN            | 94.4% (84.9–98.5); 90 and 180 pg/ml    | 100% (93.4–100); 138 and 311 µg/24 h        | 0.25         |
| <b>P-MN + P-NMN/U-A + U-NA</b>       | 94.4% (84.9–98.5); 90 and 180 pg/ml    | 77.8% (65.1–86.8); 27 and 97 µg/24 h        | <b>0.02</b>  |
| P-MN * + P-NMN */U-MN + U-NMN        | 98.1% (90.2–99.9); 58 and 45–88 pg/mL  | 100% (93.4–100); 138 and 311 µg/24 h        | 0.88         |
| <b>P-MN * + P-NMN */U-A + U-NA</b>   | 98.1% (90.2–99.9); 58 and 45–88 pg/mL  | 77.8% (65.1–86.8); 27 and 97 µg/24 h        | <b>0.008</b> |
| Comparison EIA vs. HPLC              | EIA (Plasma) Specificity (95% CI); ULN | HPLC (24 h-Urine) Specificity (95% CI); ULN | p-Value      |
| <b>P-MN + P-NMN/U-MN + U-NMN</b>     | 100% (99.6–100); 90 and 180 pg/mL      | 73.6% (70.6–76.2); 138 and 311 µg/24 h      | <b>0.001</b> |
| P-MN + P-NMN/U-A + U-NA              | 100% (99.6–100); 90 and 180 pg/mL      | 99.8% (99.2–100); 27 and 97 µg/24 h         | 0.36         |
| <b>P-MN * + P-NMN */U-MN + U-NMN</b> | 94.8% (93.2–96.1); 58 and 45–88 pg/mL  | 73.6% (70.6–76.2); 138 and 311 µg/24 h      | <b>0.004</b> |
| <b>P-MN * + P-NMN */U-A + U-NA</b>   | 94.8% (93.2–96.1); 58 and 45–88 pg/mL  | 99.8% (99.2–100); 27 and 97 µg/24 h         | <b>0.01</b>  |

Higher sensitivities and specificities are depicted in bold. The asterisks denote the corrected ULN [17].

### 3.3. ROC Curves and Comparison of AUCs

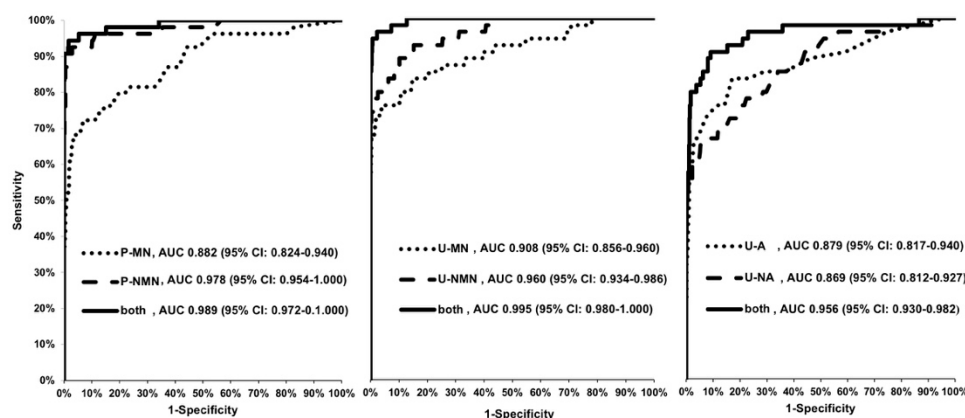
The significance level  $p$  (area = 0.5) was equal to 0.0001 for all analytes alone and in pairs, indicating that all metabolites, determined by either method, were able to correctly classify patients with and without PPGL.

When single parameters were considered, P-NMN provided the largest AUC (0.978, 95% CI: 0.954–1.000) and was superior ( $p = 0.02, 0.004,$  and  $<0.0001,$  respectively) to U-MN (0.908, 95% CI: 0.856–0.960), to U-A (0.879, 95% CI: 0.817–0.940), and to U-NA (0.869, 95% CI: 0.812–0.927), respectively, but not better ( $p = n.s.$ ) than U-NMN (0.960, 95% CI: 0.934–0.986). The AUC of P-MN (0.882, 95% CI: 0.824–0.940) was smaller ( $p = 0.02$ ) than that of U-NMN, but not different ( $p = n.s.$ ) to the other 24 h-urinary analytes (Figure 4, Table 5).

When the combined models were considered, both the AUC of the pairs of P-MN and P-NMN (0.989, 95% CI: 0.972–1.000), and of U-MN and U-NMN (0.995, 95% CI: 0.980–1.000), were larger ( $p < 0.05$ ) than that of the pair of U-A and U-NA (0.956, 95% CI: 0.930–0.975), but not different ( $p = n.s.$ ) when compared to each other (Figure 4, Table 5).

### 3.4. Post-Test Probability of PPGL

Given the 5.7% pre-test probability of PPGL in our study population (the 1.6% pre-TP based on the database of lab results of 3373 patients provided by the lab is not considered), the post-TP of a patient suffering from PPGL given a positive and negative test result of the pairs of analytes were 100% and 0% using P-MN and P-NMN together, 53% and 0% with the combined P-MN and P-NMN using the corrected ULN [17], 18% and 0% for the combined U-MN and U-NMN, and 96% and 1% for the combination of U-A and U-NA, respectively.



**Figure 4.** Area under the curve (AUC) of receiver–operator curve (ROC) of single parameters and of the combined model (pairs of metabolites).

**Table 5.** Comparisons of the AUC of ROC curves (of single parameters above, of combined parameters below).

| EIA (Plasma)/HPLC (24 h-Urine) | EIA (Plasma) AUC (95% CI)  | HPLC (24 h-Urine) AUC (95% CI) | z-Value | p-Value           |
|--------------------------------|----------------------------|--------------------------------|---------|-------------------|
| P-MN/U-MN                      | 0.882 (0.824–0.940)        | 0.908 (0.856–0.960)            | 0.64    | 0.52              |
| P-MN/U-NMN                     | 0.882 (0.824–0.940)        | <b>0.960 (0.934–0.986)</b>     | 2.39    | <b>0.02</b>       |
| P-MN/U-A                       | 0.882 (0.824–0.940)        | 0.879 (0.817–0.940)            | 0.07    | 0.94              |
| P-MN/U-NA                      | 0.882 (0.824–0.940)        | 0.869 (0.812–0.927)            | 0.31    | 0.76              |
| P-NMN/U-NMN                    | 0.978 (0.954–1.000)        | 0.960 (0.934–0.986)            | 1.02    | 0.31              |
| P-NMN/U-MN                     | <b>0.978 (0.954–1.000)</b> | 0.908 (0.856–0.960)            | 2.37    | <b>0.02</b>       |
| P-NMN/U-A                      | <b>0.978 (0.954–1.000)</b> | 0.879 (0.817–0.940)            | 2.90    | <b>0.004</b>      |
| P-NMN/U-NA                     | <b>0.978 (0.954–1.000)</b> | 0.869 (0.812–0.927)            | 3.47    | <b>&lt;0.0001</b> |
| EIA (Plasma)/HPLC (24 h-Urine) | EIA (Plasma)AUC (95% CI)   | HPLC (24 h-Urine) AUC (95% CI) | z-Value | p-Value           |
| P-MNs/U-MNs                    | 0.989 (0.972–1.000)        | 0.995 (0.980–1.000)            | 1.02    | 0.31              |
| P-MNs/U-CATs                   | <b>0.989 (0.972–1.000)</b> | 0.956 (0.930–0.975)            | 1.97    | <b>0.04</b>       |
| U-MNs/U-CATs                   | <b>0.995 (0.980–1.000)</b> | 0.956 (0.930–0.982)            | 2.12    | <b>0.03</b>       |

Higher AUCs are depicted in bold.

When analyzed with the higher ULN stipulated by the manufacturer of the 24 h-urinary metanephrines after February 2019, post-TP of PPGL given a positive test result increased to 76%, while that of a negative result remained the same. More details including summaries of sensitivities and specificities as well as calculations of post-TP of positive and negative test results based on the data of other studies are given in the Supplementary Figures.

#### 4. Discussion

To the best of our knowledge, this is the largest study comparing the diagnostic efficacy of an immunometric method for the measurement of plasma-free metanephrines (P-MNs) to that of other methods in the biochemical diagnosis of PPGL, and the only series that reports superiority of P-MNs over another method (24-h catecholamines, U-CATs). LC-MS/MS for the measurement of P-MNs or 24-h urinary metanephrines (U-MNs) have been recommended as one of the first-line tests in the biochemical diagnosis of PPGL [15]. Their availability is likely to be restricted to centers that have developed and optimized their in-house analyses, however. As outlined in the introduction, there is no lab offering a LC-MS/MS method for analysis of P-MNs in Austria. The EIAs of LDN on the other hand, developed in 2004 [27] and refined from one version to the next, are certified by ISO 9001 and ISO 13485, respectively, and satisfy the economic and medico-legal needs of public state hospitals such as the Vienna General Hospital, but still take more than 24 h to complete, probably limiting the introduction into routine use in smaller district hospitals. It may therefore be of great interest to all analysts in countries, where LC-MS/MS is not available, to know how the EIA for quantification of P-MNs perform compared to urinary measures.

Our study provides evidence of equal diagnostic performance as quantified by the AUC of ROC curves of a commercially available EIA kit for the measurement of P-MNs, when compared to a HPLC for U-MNs, and that both methods perform better than the determination of 24 h-urinary catecholamines (U-CATs) by HPLC. The use of downward corrected ULN of the former, as proposed by others [17], resulted in a slight improvement of sensitivity, at the cost of a decrease in specificity to the extent to become lower than that of both, U-MNs and U-CATs. We could not reproduce the small sensitivity of 74.1% of the EIAs, that has led others to suggest using a mathematical model for downward adjustments of the ULN. The lack of benefit of such an adjustment of ULN may have been due to the relatively small sample size of patients with PPGL, yet the number of PPGL in our study was identical to that of the German study [17]. Whether a larger number of PPGL would result in a significantly increased sensitivity remains to be determined.

The post-test probability (post-TP) that an individual is affected by a PPGL, given a positive or negative test result on the basis of the probability of the individual having the condition before the test was run (5.7%), was almost perfect and higher for the P-MNs (100% and 0%) than for U-CATs (96% and 1%), and for U-MNs (18% and 0%), respectively. Obviously, the ULN of the U-MNs (138  $\mu\text{g}/24\text{ h}$ ) and of U-NMN (311  $\mu\text{g}/24\text{ h}$ ) was too low, and has been changed by the manufacturer to 320  $\mu\text{g}/24\text{ h}$  (U-MN) and to 390  $\mu\text{g}/24\text{ h}$  (U-NMN) in February 2019, long after the study has been completed. Applying these higher ULN would have increased post-TP of PPGL given a positive test result from 18 to 76%, without changing the nearly perfect exclusion of PPGL. Given the potentially different results obtained with the new assay version after February 2019, this remains to be determined, however. Increasing ULN would increase specificity of the tests but be followed by an unacceptable decrease of sensitivity. Subjecting only patients with a high pre-test probability of PPGL to biochemical testing could lead to a better specificity. However, studies have not shown consistently better specificity (Supplementary Figures S2 and S3) with such an approach, which in routine patient care does not seem practical either.

Initial proficiency studies have suggested that P-NMN measured by the EIA of LDN was systematically lower than determined by HPLC, even after correction for the D-enantiomer of spiked samples and was ascribed to matrix effects [28]. Further quality control studies have found that the linearity of the assay for P-MN exceeded the manufacturer's claim of precision (that of P-NMN was in accord), and that both P-MN and P-NMN underestimated the Royal College of Pathologists Australasia Quality Assurance Program target value by 15.6% and 18.3%, respectively [29]. It was concluded, that the large negative bias (32%) of P-NMN measured by EIA versus by LC-MS/MS was not a matrix problem but instead related to the assay calibration [30]. Immunoassays target the natural L-enantiomer of P-MN, but calibration of the assay uses a racemic mixture of D- and L-enantiomers, because there is no commercially available calibrator material containing the L-enantiomer only [27]. Recently, measurements of P-MN by the EIA of LDN was 11% lower than those obtained by LC-MS/MS in a multicenter study of 341 patients resulting in a low sensitivity of 74.1% (14 of 54 PPGL false negative). Following the use of a recalculated lower ULN to correct for the negative bias of the EIAs, sensitivity increased to 96.2%, with only a small loss in specificity (99.3% to 95.1%). The authors conclude that until the calibration problem is corrected, ULN issued by the manufacturer should be adjusted to overcome this drawback [17]. We could not confirm these findings, suggesting that good patient care is possible using ULN stipulated by the manufacturers of the EIAs without corrections. It may be speculated that the manufacturer has resolved the calibration problem. LDN uses standards for the calibration produced from commercially available material containing racemic D-/L-MN and D-/L-NMN at a 1:1 ratio. For calibration of the assay, while the antibody used for MN detects the L-enantiomer only, that for NMN also detects the D- enantiomer at a rate of 80%. These binding properties are then used to calculate the sample weight for calibration: 1 mg MN = 0.5 mg and 1 mg NMN = 0.9 mg (personal communication, Klemm R).

The upright position leads to an increase of sympathetic output. As compared to the supine position, both plasma-free MNs (P-NMN more than P-MN) as measured by LC-MS/MS obtained from the patient in the sitting position have been suggested to result in an up to three-fold increase of false positives [31,32], so the 2014 Endocrine Society Clinical Practice Guidelines recommend that patients should be fully recumbent for at least 30 min before sampling [15]. One may assume a negligible influence of posture on the usually already highly elevated plasma concentrations of free metanephrines in patients with PPGL. However, drawing blood in the sitting position may be clinically important in those without the disease. The 25% increase of P-NMN [33] and the 27% and 30% increase of P-MN and P-NMN, respectively [31], as a result of change in posture, have been confirmed by some [34] but not all [35] authors. It has been acknowledged that blood sampling for P-MNs is frequently performed with the patient in the upright position during clinical routine work due to time constraints [36]. It is of note that ULN stipulated by the manufacturer of the EIAs of LDN were established from blood samples taken in the sitting position of 96 healthy volunteers; the first circumstance (non-recumbency) may result in falsely high and the latter (healthy individuals) in presumably falsely low ULN when

results are to be interpreted in typical screening populations of hypertensive patients. Blood sampling in our study was in the supine position in about two-thirds before and in all patients after February 2013. This may have balanced out the systematic error potentially introduced, if one assumes falsely high ULN of P-MNs established from the non-recumbent position. The degree of deviation below the ULN of false negatives was similar among all pairs of analytes (Figure 3), reassuring that the benefit of the high sensitivity was not diminished by a large range of concentrations below the ULN of either metabolite. On the other hand, the range of false positives was higher with the combined 24 h-urinary compared to the free plasma MNs (to almost 5-fold vs. up to a maximum of 3.3-fold above the ULN, respectively), suggesting that positive findings of U-MNs are less reassuring than of P-MNs, but may have been the result of the obviously falsely low ULN of the former pair of metabolites.

In comparing P-MNs by LC-ED vs. the combination of U-MNs by spectrophotometry and of U-CATs by LC-ED, it has been suggested that P-MNs may be the test of choice in high-risk patients (with familial syndromes). However, in the more common clinical situation to screen for sporadic PPGL, especially in older hypertensive patients, measurement of U-MNs in combination with U-CATs may provide comparable sensitivity with a lower rate of false-positive tests. The authors acknowledge that defining the absence of PPGL by the clinicians' evaluation or by an alternative diagnosis only was one of the limitations [37]. In our study, specificity of measurements of plasma-free MNs exceeded that of the 24 urinary-fractionated MNs and does not support their contention. Our findings are in line with some [17,38–42] but not all [30,43,44] authors assessing the performance of P-MNs measured by EIA, and is not related to sample size nor to pre-TP of PPGL of the study populations (Supplementary Figures S2 and S4, respectively). Supplementary Figure S3 gives a summary of the diagnostic performance of the EIAs of the available literature for the quantification of plasma-free metanephrines.

PPGL may relapse at a rate of 1 in 100 patient years, or with a probability as high as 25% over 5 years in patients <20 years harboring PGL >15 cm at first diagnosis [45], and may metastasize even 15 years later in patients initially thought to harbor a benign tumor [46]. There are patients with biochemical results above the ULN, but without radiological evidence of pheochromocytoma/paraganglioma (PPGL). The clinical judgement here is either a false positive biochemical result of a person free of the tumor, or a correct positive due to a PPGL too small to be detected by even the most sensitive radiological method. In both cases, long-term follow-up (FU) appears to be mandatory. Still, studies not always provide such information (Supplementary Figure S4). Given a large enough sample studied, it has been assumed, that a study's potentially incorrect categorization of patients would not make a significant difference as to sensitivity and specificity, and that the conclusion of the study would not be different [47]. Patients may develop the disease during FU, and may not have suffered from PPGL initially, even if the first lab tests have been (mildly) positive. However, PPGL not discovered are potentially life-threatening [12,13,48], and suspicion of PPGL given a positive biochemical test result may always be prudent. Six of the pertinent studies assessing P-MNs by EIA in comparison to other methods do not mention a FU of their patients defined free of PPGL [30,40,42–44], or do not report the duration of FU [38]. Three other series provide clinical, or clinical and radiological FU for 1–2 years [17,38,41] which is in line with our approach (Supplementary Figure S4). One way of interpreting concentrations above the ULN could be to not categorize test results as either "positive" or "negative", but instead as a spectrum from "PPGL likely" to "PPGL unlikely". It may be remembered in that context, that lab reference intervals usually contain 95% of the healthy population. Thus, 2.5% of healthy persons by definition have concentrations of biochemical parameters slightly above the ULN. Presenting as patients, it may not be possible to differentiate between those with and without PPGL. However, imaging and the clinical impression may help to decide, whether a positive test result is in fact ill-advised or not [48,49]. This "clinical reasoning" [50] may—in addition to the decreasing patient adherence to FU over time—provide an explanation why in clinical practice not every "positive" finding results in further diagnostic procedures. Sufficiently large study samples may be the best way to avoid bias, however [51]. The larger the study population, the smaller the 95% CI of reported sensitivities and specificities, the more credible the conclusions drawn. This applies to our study and to

all series with patients considered free of the disease by far exceeding those with PPGL (Supplementary Figure S4). Only a small fraction (8%) of patients with  $\geq 1$  positive measurements of any parameter had surgical confirmation of diagnosis (Figure 1). However, our study population was the largest of the literature investigating the performance of plasma free metanephrines by EIA for the diagnosis of PPGL (Supplementary Figure S4), and there is only one study [17] presenting data of an equal number of patients with patho-histological confirmation of diagnosis. The latter was a multicenter study, while our patients were all from a single hospital.

There are several limitations to our study. We have performed a retrospective analysis with a limited number of patients with PPGL, which may have compromised power to detect sensitivity difference. However, the overall prevalence of PPGL is very low, and a prospective study would need a very large sample population and is not available as yet. In our study, all patients were well characterized, and the sample size was among the largest of the studies that assessed the performance of an EIA for P-MNs. Also, our study was conducted in one center, potentially limiting generalizability, although it is of note that we presented data of a commercially available EIA kit widely used. We did not measure 3-methoxytyramine, the Catechol-O-Methyltransferase metabolite of dopamine, the “sibling” of MN and NMN, that recently has become available and included in the new kits of the EIA. Thus, purely dopamine-secreting PPGLs could not be detected. Moreover, drug history was not available, so specificity may have been overestimated. Interference by medications has been studied for measurement of P-MNs by LC-MS/MS, but there are no data with EIA. Incomplete collection of 24 h-urine could have resulted in falsely low U-MNs and U-CATs measurements. The use of an inappropriately low ULN of 24 h-urinary fractionated metanephrines may have compromised routine patient care, unless other diagnostic criteria have been considered in the treatment decisions. With use of the increased ULN issued by the manufacturer after February 2019, post-TP of PPGL given a positive test of the combined U-MNs improved but was still significantly below that of the pair of P-MNs. Also, we did not study P-MNs by HPLC-MS/MS, so no inference can be drawn as to that method.

One of the strengths of our study is that samples were provided by the central lab of the largest hospital in our country serving a population of more than 2 million. Additionally, all three pairs of metabolites (P-MNs, U-MNs, U-CATs) were available in every patient and at the same time within a broad spectrum of almost 1000 well-characterized patients suffering from refractory hypertension, anxiety disorders, or adrenal incidentaloma, which may be confused with PPGL, rendering the study sample representative of a typical population treated by endocrinologists including hereditary syndromes. Finally, for the first time, data were compared with use of a recently proposed ULN to adjust for the assumed systematic negative bias of the EIAs for P-MNs in addition to that of ULN stipulated by the manufacturer [17].

## 5. Conclusions

In conclusion, we have shown that the determination of P-MNs with a commercially available EIA kit is as sensitive and more specific when compared to U-MNs by HPLC, and that both methods perform better than U-CATs, the latter resulting in a low sensitivity despite excellent specificity. In the absence of HPLC-MS/MS available, state-of-the-art care of patients suspected to harbor PPGL may thus be possible with the EIAs studied. Post-test probabilities of PPGL given a positive and negative test result of P-MNs were almost 100% and 0%, respectively, with P-MNs, but only 18% and 0% with U-MNs. The systematic negative bias of P-MNs resulting in a low sensitivity by EIAs as reported previously [17] could not be reproduced, suggesting that the downward correction of ULN, as proposed by others, when measuring P-MNs by EIA [17], are not essential.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2077-0383/9/10/3108/s1>, Figure S1: Concentrations of biochemical parameters (in percent of the respective ULN) of patients with PGL ( $n = 5$ ) and PHEO ( $n = 49$ ), Figure S2: Post-test probability of PPGL given a positive and negative test result of studies summarized in Figure S4 on the basis of the respective pre-test probability, Figure S3: ULN and diagnostic sensitivities and specificities of studies comparing plasma free metanephrines (P-MNs) by immunoassays to 24 h-urinary MNs (U-MNs) and 24 h-urinary catecholamines (U-CATs), Figure S4: Sample size, indications for biochemical testing and criteria for exclusion of PPGL of the available literature on P-MNs by RIA and EIA,

Table S1: Comparisons of sensitivity (above) and specificity (below) of the single plasma compared to the single 24 h-urinary parameters using the ULN as summarized in Table 1.

**Author Contributions:** Conceptualization, W.R.; methodology, W.R. and H.K.; software, W.R. and H.K.; validation, W.R., M.K., R.M., and M.N.; formal analysis, W.R. and H.K.; data curation, H.K., W.R., M.B.N., and R.M.; writing—original draft preparation, H.K. (as part of his M.D. thesis), W.R. (above paper); writing—review and editing, W.R., H.K., M.K., M.B.N., R.M., C.S., A.K.-W., and A.K.; visualization, W.R., H.K., and M.B.N.; supervision, W.R.; project administration, W.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Pacak, K.; Tella, S.H. Pheochromocytoma and Paraganglioma. In *Endotext [Internet]*; Feingold, K.R., Anawalt, B., Boyce, A., Chrousos, G., Dungan, K., Grossman, A., Hershman, J.M., Kaltsas, G., Koch, C., Kopp, P., Eds.; MDText.com, Inc.: South Dartmouth, MA, USA, 2000–2018; (last update: 4 January 2018). Available online: <https://www.ncbi.nlm.nih.gov/books/NBK481899/> (accessed on 7 June 2020).
2. Chen, H.; Sippel, R.S.; O’Dorisio, M.S.; Vinik, A.I.; Lloyd, R.V.; Pacak, K. The North American Neuroendocrine Tumor Society Consensus Guideline for the diagnosis and management of neuroendocrine tumors: Pheochromocytoma, Paraganglioma, and Medullary Thyroid Cancer. *Pancreas* **2010**, *39*, 775–783. [[CrossRef](#)]
3. 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.* **2004**, *27*, 193–202. [[CrossRef](#)] [[PubMed](#)]
4. Brain, K.L.; Kay, J.; Shine, B. Measurement of urinary metanephrines to screen for pheochromocytoma in an unselected hospital referral population. *Clin. Chem.* **2006**, *52*, 2060–2064. [[CrossRef](#)] [[PubMed](#)]
5. Hernandez, F.C.; Sánchez, M.; Alvarez, A.; Díaz, J.; Pascual, R.; Pérez, M.; Tovar, I.; Martínez, P. A Five-Year Report on Experience in the Detection of Pheochromocytoma. *Clin. Biochem.* **2000**, *33*, 649–655. [[CrossRef](#)]
6. Václavík, J.; Stejskal, D.; Lacná, B.; Lazárová, M.; Jedelský, L.; Kadalová, L.; Janosová, M.; Frysák, Z.; Vlcek, P. Free plasma metanephrines as a screening test for pheochromocytoma in low-risk patients. *J. Hypertens.* **2007**, *25*, 1427–1431.
7. Fassnacht, M.; Arlt, W.; Bancos, I.; Dralle, H.; Newell-Price, J.; Sahdev, A.; Tabarin, A.; Terzolo, M.; Tsagarakis, S.; Dekkers, O.M. Management of adrenal incidentalomas: European Society of Endocrinology Clinical Practice Guideline in collaboration with the European Network for the Study of Adrenal Tumors. *Eur. J. Endocrinol.* **2016**, *175*, G1–G34. [[CrossRef](#)]
8. Amar, L.; Fassnacht, M.; Gimenez-Roqueplo, A.-P.; Januszewicz, A.; Prejbisz, A.; Timmers, H.; Plouin, P.-F. Long-term Postoperative Follow-up in Patients with Apparently Benign Pheochromocytoma and Paraganglioma. *Horm. Metab. Res.* **2012**, *44*, 385–389. [[CrossRef](#)]
9. Lenders, J.W.; Eisenhofer, G.; Mannelli, M.; Pacak, K. Pheochromocytoma. *Lancet* **2005**, *366*, 665–675. [[CrossRef](#)]
10. Manger, W.M.; Gifford, R.W., Jr. Pheochromocytoma: A clinical overview. In *Textbook of Hypertension*; Swales, J.D., Ed.; Blackwell Scientific Publications: Oxford, UK, 1994; pp. 941–958.
11. Manger, W.M. An Overview of Pheochromocytoma: History, Current Concepts, Vagaries, and Diagnostic Challenges. *Ann. N. Y. Acad. Sci.* **2006**, *1073*, 1–20. [[CrossRef](#)]
12. Riester, A.; Weismann, D.; Quinkler, M.; Lichtenauer, U.D.; Sommerey, S.; Halbritter, R.; Penning, R.; Spitzweg, C.; Schopohl, J.; Beuschlein, F.; et al. Life-threatening events in patients with pheochromocytoma. *Eur. J. Endocrinol.* **2015**, *173*, 757–764. [[CrossRef](#)]
13. Weismann, D.; Liu, D.; Bergen, T.; Peitzsch, M.; Raida, A.; Wagner, M.; Fassnacht, M.; Weidemann, F.; Allolio, B. Hypertension and hypertensive cardiomyopathy in patients with a relapse-free history of pheochromocytoma. *Clin. Endocrinol.* **2014**, *82*, 188–196. [[CrossRef](#)] [[PubMed](#)]
14. Guyatt, G.H.; Oxman, A.D.; Vist, G.; Kunz, R.; Falck-Ytter, Y.; Alonso-Coello, P.; Schünemann, H.J. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* **2008**, *336*, 924–926. [[CrossRef](#)] [[PubMed](#)]

15. Lenders, J.; Duh, Q.-Y.; Eisenhofer, G.; Gimenez-Roqueplo, A.-P.; Grebe, S.K.G.; Murad, M.; Naruse, M.; Pacak, K.; Young, W.F. Pheochromocytoma and Paraganglioma: An Endocrine Society Clinical Practice Guideline. *J. Clin. Endocrinol. Metab.* **2014**, *99*, 1915–1942. [CrossRef] [PubMed]
16. Därr, R.; Kuhn, M.; Bode, C.; Bornstein, S.R.; Pacak, K.; Lenders, J.W.; Eisenhofer, G. 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. *Endocrine* **2017**, *56*, 495–503. [CrossRef] [PubMed]
17. Weismann, D.; Peitzsch, M.; Raida, A.; Prejbisz, A.; Gosk, M.; Riester, A.; Willenberg, H.S.; Klemm, R.; Manz, G.; Deutschbein, T.; et al. Measurements of plasma metanephrines by immunoassay vs liquid chromatography with tandem mass spectrometry for diagnosis of pheochromocytoma. *Eur. J. Endocrinol.* **2015**, *172*, 251–260. [CrossRef]
18. Labor Diagnostika Nord, Nordhorn, Germany (LDN). Instruction for Use. 2-MET Plasma ELISA Fast Track. Available online: <https://ldn.de/product/2-met-plasma-elisa-fast-track/> (accessed on 7 June 2020).
19. Pussard, E.; Neveux, M.; Guigueno, N. Reference intervals for urinary catecholamines and metabolites from birth to adulthood. *Clin. Biochem.* **2009**, *42*, 536–539. [CrossRef]
20. Dawson-Saunders, B.; Trapp, R.G. Evaluating diagnostic procedures. In *Basic and Clinical Biostatistics*, 2nd ed.; Appleton & Lange: New York City, NY, USA, 1994; pp. 232–248.
21. Wilson, E.B. Probable inference, the law of succession, and statistical inference. *J. Am. Stat. Ass.* **1927**, *22*, 209–212. [CrossRef]
22. Newcombe, R.G. Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Stat. Med.* **1998**, *17*, 857–872. [CrossRef]
23. Hanley, J.A.; McNeil, B.J. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* **1983**, *148*, 839–843. [CrossRef]
24. Fagan, T.J. Nomogram for Bayes's Theorem. *N. Engl. J. Med.* **1975**, *293*, 257.
25. Caraguel, C.G.; Vanderstichel, R. The two-step Fagan's nomogram: Ad hoc interpretation of a diagnostic test result without calculation. *Evid.-Based Med.* **2013**, *18*, 125–128. [CrossRef] [PubMed]
26. Schwartz, A. Diagnostic Test Calculator-Version 2010042101: Available Under the Clarified Artistic License. 2006. Available online: <http://araw.mede.uic.edu/cgi-bin/testcalc.pl> (accessed on 7 June 2020).
27. Manz, B.; Kuper, M.; Booltink, E.; Fischer-Brügge, U. Development of Enantioselective Immunoassays for Free Plasma Metanephrines. *Ann. N. Y. Acad. Sci.* **2004**, *1018*, 582–587. [CrossRef] [PubMed]
28. Pillai, D.; Ross, H.A.; Kratzsch, J.; Pedrosa, W.; Kema, I.; Hoad, K.; Rouaix, N.; Fathi, M.; Nader, H.; Mathian, B.; et al. Proficiency test of plasma free and total metanephrines: Report from a study group. *Clin. Chem. Lab. Med.* **2009**, *47*, 786–790. [CrossRef]
29. Horowitz, G.L.; Boyd, J.C.; Ceriotti, F.; Garg, U.; Horn, P.; Pesce, A.; Sine, H.E.; Zakowski, J. *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory: Approved Guideline*, 3rd ed.; Clinical and Laboratory Standards Institute (CLSI): Wayne, PA, USA, 2008; CLSI document c28-a3; ISBN 1-56238-682-4.
30. 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.* **2012**, *50*, 105–110. [CrossRef]
31. Lenders, J.; Willemsen, J.J.; Eisenhofer, G.; Ross, H.A.; Pacak, K.; Timmers, H.J.L.M.; Sweep, F.C.G.J. Is Supine Rest Necessary before Blood Sampling for Plasma Metanephrines? *Clin. Chem.* **2007**, *53*, 352–354. [CrossRef]
32. Därr, R.; Pamporaki, C.; Peitzsch, M.; Miehle, K.; Prejbisz, A.; Peczkowska, M.; Weismann, D.; Beuschlein, F.; Sinnott, R.; Bornstein, S.R.; et al. Biochemical diagnosis of phaeochromocytoma using plasma-free normetanephrine, metanephrine and methoxytyramine: Importance of supine sampling under fasting conditions. *Clin. Endocrinol.* **2013**, *80*, 478–486. [CrossRef] [PubMed]
33. Eisenhofer, G. Editorial: Biochemical diagnosis of pheochromocytoma—Is it time to switch to plasma free metanephrines? *J. Clin. Endocrinol. Metab.* **2003**, *88*, 550–552. [CrossRef] [PubMed]
34. Deutschbein, T.; Unger, N.; Jaeger, A.; Broecker-Preuss, M.; Mann, K.; Petersen, S. Influence of various confounding variables and storage conditions on metanephrine and normetanephrine levels in plasma. *Clin. Endocrinol.* **2010**, *73*, 153–160.
35. Raber, W.; Raffesberg, W.; Bischof, M.; Scheuba, C.; Niederle, B.; Gasic, S.; Waldhäusl, W.; Roden, M. Diagnostic efficacy of unconjugated plasma metanephrines for the detection of pheochromocytoma. *Arch. Intern. Med.* **2000**, *160*, 2957–2963. [CrossRef]

36. Eisenhofer, G.; Lattke, P.; Herberg, M.; Siegert, G.; Qin, N.; Därr, R.; Hoyer, J.; Villringer, A.; Prejbisz, A.; Januszewicz, A.; et al. Reference intervals for plasma free metanephrines with an age adjustment for normetanephrine for optimized laboratory testing of pheochromocytoma. *Ann. Clin. Biochem. Int. J. Lab. Med.* **2012**, *50*, 62–69. [[CrossRef](#)]
37. Sawka, A.M.; Jaeschke, R.; Singh, R.J.; Young, W.F., 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*, 553–558. [[CrossRef](#)] [[PubMed](#)]
38. Procopiou, M.; Finney, H.; Akker, S.A.; Chew, S.L.; Drake, W.M.; Burrin, J.; Grossman, A.B. Evaluation of an enzyme immunoassay for plasma-free metanephrines in the diagnosis of catecholamine-secreting tumors. *Eur. J. Endocrinol.* **2009**, *161*, 131–140. [[CrossRef](#)] [[PubMed](#)]
39. Christensen, T.T.; Frystyk, J.; Poulsen, P.L. Comparison of plasma metanephrines measured by a commercial immunoassay and urinary catecholamines in the diagnosis of pheochromocytoma. *Scand. J. Clin. Lab. Investig.* **2011**, *71*, 695–700. [[CrossRef](#)] [[PubMed](#)]
40. Tanaka, Y.; Isobe, K.; Ma, E.; Imai, T.; Kikumori, T.; Matsuda, T.; Maeda, Y.; Sakurai, A.; Midorikawa, S.; Hataya, Y.; et al. Plasma free metanephrines in the diagnosis of pheochromocytoma: Diagnostic accuracy and strategies for Japanese patients. *Endocr. J.* **2014**, *61*, 667–673. [[CrossRef](#)]
41. Pussard, E.; Chaouch, A.; Said, T. Radioimmunoassay of free plasma metanephrines for the diagnosis of catecholamine-producing tumors. *Clin. Chem. Lab. Med.* **2014**, *52*, 437–444. [[CrossRef](#)]
42. Bílek, R.; Zelinka, T.; Vlček, P.; Dušková, J.; Michalský, D.; Novák, K.; Václavíková, E.; Widimský, J. Radioimmunoassay of Chromogranin A and Free Metanephrines in Diagnosis of Pheochromocytoma. *Physiol. Res.* **2017**, *66* (Suppl. 3), 397–408.
43. Gao, Y.-C.; Lu, H.-K.; Luo, Q.-Y.; Chen, L.-B.; Ding, Y.; Zhu, R.-S. Comparison of free plasma metanephrines enzyme immunoassay with <sup>131</sup>I-MIBG scan in diagnosis of pheochromocytoma. *Clin. Exp. Med.* **2008**, *8*, 87–91. [[CrossRef](#)]
44. Unger, N.; Pitt, C.; Schmidt, I.L.; Walz, M.K.; Schmid, K.W.; Philipp, T.; Mann, K.; Petersenn, S. Diagnostic value of various biochemical parameters for the diagnosis of pheochromocytoma in patients with adrenal mass. *Eur. J. Endocrinol.* **2006**, *154*, 409–417. [[CrossRef](#)]
45. Plouin, P.F.; Amar, L.; Dekkers, O.M.; Fassnacht, M.; Gimenez-Roqueplo, A.-P.; Lenders, J.W.M.; Lussey-Lepoutre, C.; Steichen, O.; On behalf of the Guideline Working Group. European Society of Endocrinology Clinical Practice Guideline for long-term follow-up of patients operated on for a pheochromocytoma or a paraganglioma. *Eur. J. Endocrinol.* **2016**, *174*, G1–G10. [[CrossRef](#)]
46. Koratala, A.; Bhattacharya, D.; Wakefield, D.N. Benign-appearing pheochromocytoma with late recurrence. *Am. J. Med. Sci.* **2017**, *353*, e13. [[CrossRef](#)]
47. Lenders, J.W.M.; Pacak, K.; Walther, M.M.; Linehan, W.M.; Mannelli, M.; Friberg, P.; Keiser, H.R.; Goldstein, D.S.; Eisenhofer, G. Biochemical Diagnosis of Pheochromocytoma. *JAMA* **2002**, *287*, 1427–1434. [[CrossRef](#)] [[PubMed](#)]
48. Pacak, K.; Eisenhofer, G.; Ahlman, H.; Bornstein, S.R.; Gimenez-Roqueplo, A.-P.; Grossman, A.B.; Kimura, N.; Mannelli, M.; McNicol, A.M.; Tischler, A.S. Pheochromocytoma: Recommendations for clinical practice from the First International Symposium. *Nat. Clin. Pract. Endocrinol. Metab.* **2007**, *3*, 92–102. [[CrossRef](#)] [[PubMed](#)]
49. Grossman, A.; Pacak, K.; Sawka, A.; Lenders, J.W.M.; Harlander, D.; Peaston, R.T.; Reznick, R.; Sisson, J.; Eisenhofer, G. Biochemical Diagnosis and Localization of Pheochromocytoma: Can We Reach a Consensus? *Ann. N. Y. Acad. Sci.* **2006**, *1073*, 332–347. [[CrossRef](#)] [[PubMed](#)]
50. Mattingly, C. What is Clinical Reasoning? *Am. J. Occup. Ther.* **1991**, *45*, 979–986. [[CrossRef](#)] [[PubMed](#)]
51. Holmes, T.H. Ten categories of statistical errors: A guide for research in endocrinology and metabolism. *Am. J. Physiol. Metab.* **2004**, *286*, E495–E501. [[CrossRef](#)]

