RESEARCH

Clinical decision making in small non-functioning VHL-related incidentalomas

Roland Därr[1,2], Jonas Kater[2], Peggy Sekula [ID][3], Birke Bausch[4], Tobias Krauss[5], Christoph Bode[1], Gerd Walz[2], Hartmut P Neumann[6] and Stefan Zschiedrich[2]

[1]Department of Cardiology and Angiology I, Heart Center Freiburg University, Faculty of Medicine, University of Freiburg, Freiburg, Germany
[2]Department of Medicine IV, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany
[3]Institute of Genetic Epidemiology, University Medical Center Freiburg, Freiburg, Germany
[4]Department of Medicine II, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany
[5]Department of Radiology, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany
[6]Section for Preventive Medicine, Faculty of Medicine, University of Freiburg, Freiburg, Germany

Correspondence should be addressed to R Därr: roland_daerr@web.de

Abstract

The optimal treatment strategy for patients with small non-functioning VHL-related incidentalomas is unclear. We searched the Freiburg VHL registry for patients with radiologic evidence of pheochromocytoma/paraganglioma (PHEO/PGL). In total, 176 patients with single, multiple, and recurrent tumours were identified (1.84 tumours/patient, range 1–8). Mean age at diagnosis was 32 ± 16 years. Seventy-four percent of tumours were localised to the adrenals. Mean tumour diameter was 2.42 ± 2.27 cm, 46% were <1.5 cm. 24% of tumours were biochemically inactive. Inactive tumours were significantly smaller than active PHEO/PGL at diagnosis (4.16 ± 2.80 cm vs 1.43 ± 0.45 cm; *P* < 0.025) and before surgery (4.89 ± 3.47 cm vs 1.36 ± 0.43 cm; *P* < 0.02). Disease was stable in 67% of 21 patients with evaluable tumours ≤1.5 cm according to RECIST and progressed in 7. Time till surgery in these patients was 29.5 ± 20.0 months. A total of 155 patients underwent surgery. PHEO/PGL was histologically excluded in 4 and proven in 151. Of these, one had additional metastatic disease, one harboured another tumour of a different type, and in 2 a second surgery for suspected disease recurrence did not confirm PHEO/PGL. Logistic regression analysis revealed 50% probability for a positive/negative biochemical test result at 1.8 cm tumour diameter. Values of a novel symptom score were positively correlated with tumour size ($R_s = 0.46$, *P* < 0.0001) and together with a positive biochemistry a linear size predictor (*P* < 0.01). Results support standardised clinical assessment and measurement of tumour size and metanephrines in VHL patients with non-functioning incidentalomas <1.5 cm at one year following diagnosis and at individualised intervals thereafter depending on evolving growth dynamics, secretory activity and symptomatology.

Introduction

Von Hippel-Lindau (VHL) disease is an inherited autosomal-dominant tumour predisposition syndrome caused by inactivating mutations of the *VHL* tumour suppressor gene. The disorder is characterised by the frequent development of CNS and retinal hemangioblastomas, renal clear cell carcinomas, and neuroendocrine tumours. Furthermore, endolymphatic sac tumours, cystadenomas, and cysts occur in various visceral organs (1, 2).

Pheochromocytomas (PHEO) and paragangliomas (PGL) are found in up to 30% of patients with VHL disease

depending on the underlying mutation (3, 4, 5, 6). These tumours are often bilateral, multifocal, and recurrent (7) and are also referred to as cluster-1 PHEO/PGL (8, 9, 10). They mainly produce norepinephrine owing to a deficiency of the enzyme phenylethanolamine-N-methyltransferase (11), and at the gene expression level they show increased hypoxic signalling (10, 12, 13, 14, 15). Of note, the risk of malignant transformation is remarkably low in *VHL*-mutated PHEO/PGL usually not exceeding 5% in published series compared with 30–60% of most other PHEO/PGL with cluster-1 identity (3, 6, 16, 17, 18, 19, 20).

The advent of high-resolution CT and MRI and their widespread use in routine surveillance programs has considerably helped early detection of serious and potentially life-threatening VHL organ manifestations, and prompt treatment has since clearly expanded the life expectancy of afflicted patients and their kindred (21). Accompanying this gratifying trend has been a noticeable increase in the discovery of asymptomatic adrenal and extra-adrenal masses suspicious of PHEO/PGL (22, 23, 24). These so-called incidentalomas tend to be smaller than sporadic PHEO/PGL, may exhibit a remarkably slow growth rate, and do not release excess catecholamines in a substantial proportion of cases (3, 6, 25). Obviously, therefore, the decision for surgery and histological verification of PHEO/PGL in patients with these tumours may be a difficult undertaking. The situation is further complicated by the fact that current VHL surveillance protocols do not address the issue of incidentalomas, and in particular do not provide recommendations to guide timing and scope of follow-up tests in patients where biochemical evidence of PHEO/PGL cannot be obtained (26, 27, 28, 29, 30). In the present study, we took advantage of a comprehensive documentation of patients with VHL disease and radiologic evidence of PHEO/PGL, the so-called Freiburg VHL registry, in order to find out whether readily available information on fundamental tumour qualities, that is size, secretory activity, growth dynamics and symptoms, could serve as a basis for a both safe and efficient treatment strategy in patients with small non-functioning VHL-related incidentalomas.

Materials and methods

Definitions

The term 'non-functioning' was used as a synonym for 'biochemically inactive' or 'biochemically silent' to define PHEO/PGL without laboratory evidence of excess hormone secretion. The terms 'symptomatic' and 'asymptomatic' (clinically 'silent') were, respectively, applied as meaning PHEO/PGL with or without disease-related clinical signs or symptoms according to the judgment of the investigator at the time of first diagnosis.

Study design and patients

For the present retrospective observational study the VHL registry of the University Medical centre Freiburg (the Freiburg VHL registry), Germany, was searched for all patients with CT/MRI results suspicious of PHEO/PGL. Relevant demographic and clinical information on these patients, in particular gender, age at diagnosis, manifesting symptoms, type of mutation, tumour location and tumour size, biochemical test results, time of surgery, histopathology findings, time and results of follow-up visits and patient outcomes was retrieved from the respective electronic files as well as supplementary databases, and the collected data were then anonymised so that the individual was no longer identifiable. Use of the anonymised data for further analysis was approved by the local review board and ethics committee of Freiburg University (EK-FR 260/18). Written informed consent had been given by all patients.

Sign and symptom score

A clinical evaluation score (CES) was developed in order to allow for the uniform categorisation into symptomatic and asymptomatic patients. To that end, patients were ranged according to the documented presence/absence of the common signs and symptoms of PHEO/PGL in descending order of overall observed percentages, where signs comprised hypertension and diabetes mellitus, and symptoms the classic triad headaches, palpitations and sweating or a single symptom of it, as well as weight loss, flushing, pallor, nausea, vertigo, tiredness, and changes in the capacity to meet the ordinary demands of life or changes in mood or perception (31, 32, 33). Symptoms not typically related to PHEO/PGL, such as sleep disturbances, diarrhoea, epistaxis or abdominal, flank or back pain were also considered and collectively classified as 'other'. From the upper and lower 25% of patients the 'item' discrimination index level (ID-Idx) and the corrected 'item'-to-scale correlation (CI/S-cor) were calculated for each sign/symptom. Correlation values equal to or greater than ≥0.37 were considered to signify symptoms with a high diagnostic weight (34). Difficulty-by-discrimination plots were used to assess the performance of individual

'items', and to corroborate the plausibility and consistency of the developed score. The combined findings were then used to establish a graded rating system consisting of 11 selected signs and symptoms.

Assessment of tumour size and tumour growth

Tumour size was estimated by unidimensional measurement of the longest tumour diameter on representative CT/MRI scans. For multiple (up to 4) tumours occurring in one patient at the same time the longest sum diameter was calculated and tumour growth was determined by a change from baseline to last follow up before surgery according to the revised version of the response evaluation criteria in solid tumours (RECIST) guideline, version 1.1. Progressive disease was defined as a relative increase in tumour size of at least 20% (35).

Statistical analysis

R-studio (Version 1.2.5001) was applied for statistical analyses. For comparison of categorical, ordinal, and continuous numerical values exact binomial test, χ^2-test or Fisher's exact test (for count data <5) were performed. Shapiro–Wilk normality test was used to assess the normal distribution of the data and log transformation was performed if indicated. Parametric (e.g. *t*-test) and non-parametric tests (e.g. Wilcoxon test, Kruskal–Wallis test) were carried out to compare groups, as appropriate. Significance levels were adjusted according to Bonferroni when required. Logistic regression models were fitted on binary outcome variables (symptoms present/absent; biochemical test result positive/negative). All models were first fitted by backward selection of a maximum model of complete cases with respect to the baseline characteristics gender, age at tumour diagnosis and underlying mutation, and single parameters of interest were then sequentially added to these models from the available cases. R-squared (R^2) was determined as a measure of the goodness-of-fit of the models. Spearman's coefficient was calculated for correlations. Receiver operating characteristic (ROC) curves were computed to assess predictive performance. Area under ROC curves (AUC) were compared based on the method of DeLong (36). Threshold analysis was carried out for determining sensitivities, specificities and corresponding score values and tumour size. CI were calculated using the 'pROC' package. Tumours (not patients) were used as the unit of analysis, when necessary, in order to allow inclusion of all relevant information but ignoring potential dependencies. Multiple tumours were treated as one large tumour by adding up the longest diameters from all the tumours of a patient (see previously), and tumour recurrence was accounted for by using the consecutive tumour count as a factor.

Results

Patient characteristics

Search of the VHL registry, relevant files and supplementary databases revealed 176 patients with imaging results suspicious of PHEO/PGL (1.84 tumours/patient, range 1–8; 239 adrenal tumours, 85 extra-adrenal tumours) out of a total of 928 patients with VHL disease (19%). Surgery was documented in 155 patients, and the diagnosis PHEO/PGL was confirmed by histology in 151 (1.42 tumours/patient; 172 adrenal tumours, 43 extra-adrenal tumours) and excluded in 4 (1.0 tumour/patient; 3 adrenal, 1 extra-adrenal). One patient with confirmed PHEO/PGL was also found to harbour another tumour with a different histology, and one had additional metastatic disease (1 extra-adrenal tumour). In 59 patients with confirmed PHEO/PGL recurrent disease was subsequently suspected from follow-up imaging. Of these 2 underwent a second surgery and PHEO/PGL was excluded in each of the cases. Thus, in total there were three groups of patients with partially significant overlap, one group consisting of 151 patients with histologically confirmed PHEO/PGL (one patient with additional metastatic disease), a second group of 78 patients with suspected PHEO/PGL (21 patients of 176 with suspected PHEO/PGL who were not brought to surgery, 57 patients of 151 with confirmed PHEO/PGL and suspected disease recurrence), and the third group of 7 patients with a different tumour entity (4 patients of 155 operated on for suspected PHEO/PGL which was not confirmed, one patient of 151 with confirmed PHEO/PGL and an additional tumour of a different type, 2 patients of 59 with confirmed PHEO/PGL and suspected disease recurrence which was excluded in a second surgery). Patients with PHEO/PGL were significantly younger than patients where PHEO/PGL was not verified (mean age 32 ± 16 years vs 52 ± 16 years, $P < 0.005$). Also, mean age at diagnosis was significantly lower (29 ± 15 years vs 36 ± 16 years; $P < 0.05$) and the incidence of PHEO/PGL was significantly higher (0.55 vs 0.43; $P < 0.05$) in men compared to women. The prevalence of PHEO/PGL in patients with tumours <1.5 cm was ~96% (Table 1).

Table 1 Patient characteristics.

	Confirmed PHEO/PGL[a] + 1 metastatic	Excluded PHEO/PGL[a]	Suspected PHEO/PGL[b]	*P*-value[j]
No.[i]	151	7	78	
Gender				0.980
female (No.)	66[k]	3	33	
male (No.)	85[k]	4	45	
Age at diagnosis (years)	32 ± 16[e,f]	52 ± 16[f]	41 ± 17[e]	<0.0001
female	36 ± 16[g]	52 ± 24	48 ± 13	
male	29 ± 15[g]	53 ± 10	35 ± 17	
Biochemical test[c] (No.) (%)	77 (51%)	5 (71%)	48 (62%)	0.6258
positive/negative (%)	70/30	20/80	42/58	<0.002
Imaging test (No.)[d] (%)	66 (44%)	6 (86%)	71 (91%)	<0.004
Tumours/patient (range)	1.43 (1–7)	1.0 (NA)	1.29 (1–8)	0.721
Tumour size (cm)	3.4 ± 2.7[h]	2.2 ± 0.9	1.49 ± 1.36[h]	<0.0001
Clinical classification[d] (No.)	81	5	67	<0.0001
asymptomatic (%)	65	80	94	
symptomatic (%)	35	20	6	

[a]Histopathology; [b]CT/MRI; [c]At initial diagnosis or subsequent visits; [d]At initial diagnosis; [e]Confirmed vs suspected PHEO/PGL *P* < 0.0002; [f]Confirmed PHEO/PGL vs excluded PHEO/PGL *P* < 0.005; [g]Age female vs male <0.05; [h]Size confirmed vs suspected PHEO/PGL <0.0001; [i]Note that the total number of patients in the three study groups does not amount to 176 due to a partially significant overlap as outlined in the 'Results' section; [j]*P*-values for the overall comparison among the three groups, pairwise comparisons: [k]female vs male *P*< 0.05; NA, not applicable; ±, means ± s.d.

Biochemical test results and tumour symptoms

Biochemical test results were documented in 51% (n=77) and 62% (n=48) of patients with confirmed or suspected PHEO/PGL, and measurement of metanephrines in plasma or urine (MNs) were reported in 30% (n=23) and 48% (n=23) of these patients. Respective information relating to tumour symptoms was available from 54% (n=81) and 86% (n=67) of patients. The proportion of symptomatic tumours in these patients was 35 and 6% as judged by the investigator at the time of first diagnosis (Table 1).

VHL gene mutations and tumour burden

Patients with confirmed and suspected PHEO/PGL harboured 53 different *VHL* gene mutations. The genetic aberration most commonly encountered was the missense 292T>C (Tyr98His) mutation at 43.2%, also referred to as 'Black Forest' (BF) mutation (5), followed by the missense mutations 499C>T (Arg167Trp), 500G>A (Arg167Gln) and 562C>G (Leu188Val) at less than 6% each. Patients with extra-adrenal tumour location were significantly more likely to carry the BF mutation compared to those with adrenal gland tumours (P < 0.0001). However, neither median tumour number (1.0 for $BF_{present}$ (range 1–8) vs 1.0 for BF_{absent} (range 1–4); P=0.1275) nor mean tumour size at diagnosis ($BF_{present}$ 2.68 ± 2.75 vs BF_{absent} 2.44 ± 1.98; P=0.367) were related to the underlying type of mutation. These results rule out enhanced tumour-promoting effects of individual mutations in patients with confirmed PHEO/PGL.

Biochemical test results, symptoms and tumour size in patients with PHEO or PGL

Comparison between PHEO and PGL in patients with either tumour entity did not reveal significant differences in the proportion of biochemically inactive and active tumours (PHEO (n=70) 0.52; PGL (n=20) 0.43; P=0.7199). Also, there were no significant differences in the proportion of symptomatic and asymptomatic tumours in these two subgroups of patients (PHEO (n=74) 0.45; PGL (n=22) 0.22; P=0.2897), and there were no significant differences in mean tumour size (PHEO (n=59): 3.25 ± 2.46 cm; PGL (n=19): 2.62 ± 1.57 cm; P=0.4912). Taken together, these results show that a high degree of similarity exists between PHEO and PGL with regard to three core tumour qualities thereby justifying the strategy adopted in the present study where PHEO and PGL were combined into one single group for statistical reasons.

Tumour size and tumour growth

CT/MRI scans from 119 patients (68%) served as the basis for assessment of tumour size at the time of initial diagnosis. Mean tumour diameter was 2.42 ± 2.27 cm (number of tumours n=177; range 0.3–14.9 cm). Nearly half of the tumours (46%) were smaller than 1.5 cm. CT/MR scans obtained prior to surgery were available from 21 of patients with small-sized tumours for evaluation of tumour growth. Seven (33%) of these 21 patients (1.43 tumours/patient) met the criteria of progressive disease (for details see 'Materials and methods' section). In these

patients mean tumour diameter at baseline was 1.47 ± 0.85 cm, mean change in size to surgery was 1.12 ± 0.59 cm, and mean time elapsed to last follow up (LFU) before surgery was 29.5 ± 20 months. There was no evidence for segregation of a particular mutation with the progressive phenotype and there was also no indication for discrete clinical, biochemical or imaging features (Table 2). Change in size to surgery did not qualify for tumour progression in 67% of patients ($n = 14$) with tumours ≤ 1.5 cm (1.1 tumours/patient; mean tumour diameter 1.13 ± 0.17 cm, $P = 0.66$; LFU 28 ± 42 months, $P = 0.16$). PHEO/PGL in these patients were designated as stable disease (Fig. 1).

Clinical evaluation score (CES)

A clinical scoring system (CES) was developed on a total of 11 signs and symptoms of PHEO/PGL enabling conversion of qualitative dichotomous outcomes to a numerical measure and stratification of patients according to scope and severity of complaints at individual level (for details see 'Materials and methods' section). In this system, a three-grade linear score (0.0, 1.0, 2.0) was assigned to each of four items (headaches, palpitations, sweating, hypertension) with a high diagnostic weight (CI/S-cor ≥ 0.37), a second differently weighted three-grade linear score (0.0, 0.5, 1.0) to another six items (tiredness, weight loss, flushing/pallor, nausea, vertigo, changes in mood/perception) with a CI/S-cor value <0.37, and a two-grade score (0.0, 1.0) to one item (diabetes mellitus). The symptom triad and symptoms collectively termed as 'other' were not considered, since neither was found to enhance the diagnostic precision of the final comprehensive scoring system representing the sum of the individual item scores (Table 3).

Discriminative properties of the CES

Documented tumour symptoms and binary classification of patients with confirmed PHEO/PGL as symptomatic or symptom-free by the investigator were used for the construction of ROC curves and assessment of threshold score values with corresponding diagnostic sensitivity/specificity (SE/SP) pairs. With this approach, a reliable distinction between symptomatic and asymptomatic patients was achieved using cut-off score values of 2.74 and 2.54 for recorded symptoms at diagnosis and before surgery, respectively. SE/SP pairs for these values were 94% (CI 84–100%)/98% (CI 95–100%) and 90% (CI 79–97%)/95% (CI 88–100%), and areas under the ROC curve (AUCs) were 0.991 (CI 0.9787–1.00) and 0.975 (CI 0.9528–0.9977) ($P = 0.2355$). Hence, it would appear that CES values equal to or smaller than 2.5 correctly define patients with asymptomatic PHEO/PGL while values equal to or greater than 3.0 indicate symptomatic disease (not shown).

Correlation between CES and tumour size

Spearman's coefficient of correlation between CES values and tumour size was $R_s = 0.46$ (<0.0001; $n = 85$)

Table 2 Characteristics of PHEO/PGL in patients with 'progressive' disease according to RECIST.

Pt.	Tm.	TNL	Loc.	F/M	MUT	A/D	Sz	CES	Sz$_{LFU}$	CES$_{LFU}$	Time	BC	Growth (%)
1	1	2/4	L	F	Pool	30	1.6	0 (A)	2.2	2 (S)	19	Neg	37.5
2	2	5/7	L	F	BF	37	1.3	1 (A)	2.8	2 (A)	28	Pos	115
3	3	1/4	R	M	BF	14	3.0	0 (A)	4.4	5 (S)	74	Pos	46.7
3	4	3/4	L	M	BF	25	1.0	0 (A)	1.6	2 (A)	15	Pos	60
3	5	4/4	PGL	M	BF	27	1.0	1 (A)	1.7	1 (A)	25	NA	70
4	6	1/3	R$_{con}$	F	Pool	26	1.0	0 (A)	2.5	2 (A)	14	Pos	150
4	7	2/3	L	F	Pool	26	1.2	0 (A)	2.6	0 (A)	39	Pos	117
5	8	2/2	R	M	BF	37	0.6	2 (A)	2.8	0 (A)	51	Neg	367
6	9	1/1	B	F	Pool	69	3.0	0 (A)	4.1	NA (S)	20	Pos	36.7
	9.1		L				1.5		2.0				33.3
	9.2		R				1.0		1.4				40
	9.3		R				0.5		0.7				40
7	10	1/1	R	F	Pool	19	1.0	4 (S)	1.2	3 (S)	10	Pos	20

A/D, age at diagnosis; A/S, disease-related symptoms absent/present as documented; B, bilateral PHEO; BC, biochemical test result positive (Pos) or negative (Neg) for the respective tumour at any point in time in the course of the disease; BF, 'Black Forest' mutation; CES, clinical evaluation score sum at diagnosis and at last follow-up (CESLFU) (asymptomatic ≤2.5, symptomatic ≥3.0, see also 'Results' section) ; F/M, female/male gender; L, left-sided pheochromocytoma (PHEO) ; Loc, tumour location; MUT, underlying mutation; NA, not applicable; PGL, paraganglioma; Pool, other mutation; Pt, patient number; R, right-sided PHEO; Rcon, right-sided conglomerate tumour (2x) ; Sz, tumour size at diagnosis and at last follow-up (SzLFU) ; Time, time elapsed to last follow-up prior to surgery (months) ; Tm, tumour number; TNL, current tumour number in patients with multiple consecutive tumours.
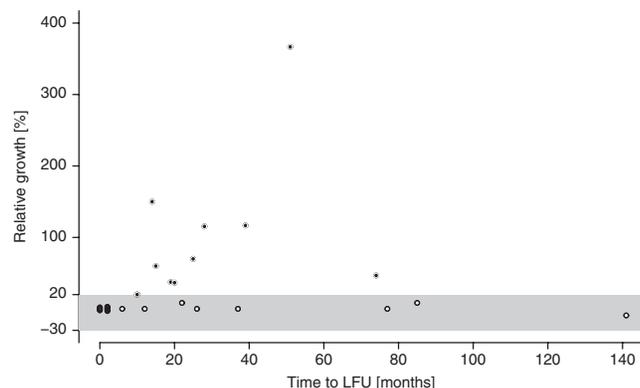
Figure 1

Relative growth of small-sized (≤1.5 cm) PHEO/PGL. Mean tumour size at baseline was 1.47 ± 0.85 cm (for details see also Table 2) vs 1.13 ± 0.17 cm ($P = 0.66$) and mean time elapsed to last follow-up prior to surgery 29.5 ± 20 months vs 28 ± 42 months ($P = 0.16$) in patients with, respectively, progressive ($n = 7$, 1.43 tumours/patient) and stable ($n = 14$, 1.10 tumours/patient) disease. Black solid circles indicate progressive growth and open circles stable tumour size. The shaded area defines the upper and lower boundaries of stable disease according to RECIST.

at baseline and $R_s = 0.37$ ($P < 0.001$; $n = 82$) prior to surgery, respectively. In addition, a parallel upward shift was observed of the respective linear regression lines for functionally active PHEO/PGL compared with biochemically silent tumours (Fig. 2). These results are in agreement with the findings described subsequently on an apparent link between PHEO/PGL size and secretory activity, thus lending credit to the validity of the proposed CES system.

Table 3 Clinical evaluation score (CES) for discrimination between symptomatic and asymptomatic tumour patients.

Symptoms	Absent	Present	
		Intermittent	Persistent
Headaches	0	1	2
Palpitations	0	1	2
Sweating	0	1	2
Tiredness	0	0.5	1
Nausea	0	0.5	1
Vertigo	0	0.5	1
Psyche[a]	0	0.5	1
Weight loss	0	0.5	1
Flushing/pallor	0	0.5	1
Hypertension	0	Chronic: 1	Paroxysmal: 2
Diabetes	0	1	

The scoring system assigns weighted values to 11 selected signs and symptoms (for details see 'Materials and methods' section). The maximum score sum value is 15. Values ≤2.5 and ≥3.0 define asymptomatic and symptomatic disease, respectively.
[a]Mental impairments, for example, anxiety, changes in mood and perception, nervousness, lethargy.
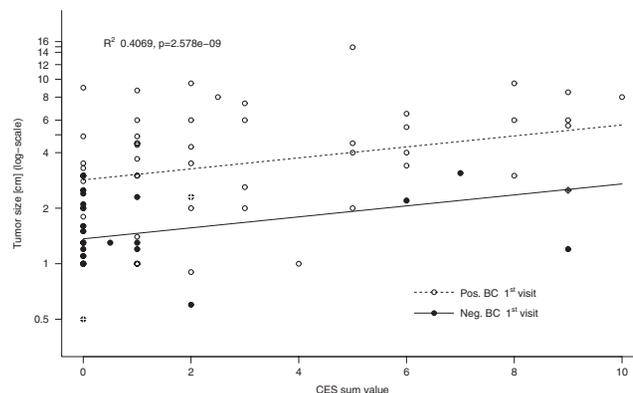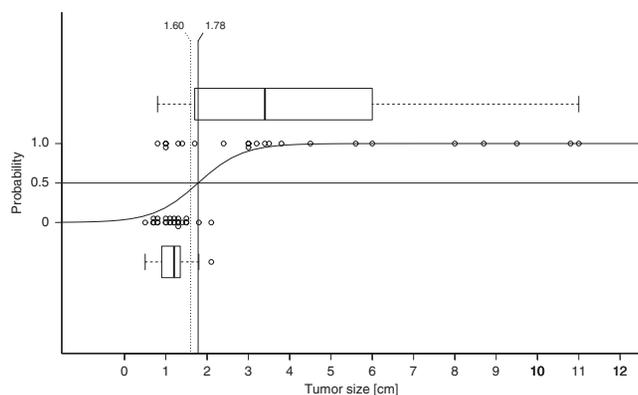

Figure 2

Clinical evaluation score (CES) sum value vs tumour size. The CES sum value at first diagnosis is positively correlated with tumour size and a linear size predictor for both biochemically active (open circles; 52 tumours) and non-functioning PHEO/PGL (solid black circles; 28 tumours). Crossed circles mark tumours where PHEO/PGL was excluded. Multiple tumours and recurrent disease were treated as outlined in the 'Materials and methods' section. Note the parallel upward shift (identical R^2) of the equation regression lines for biochemically active (dashed line) compared to non-functioning (solid line) tumours indicating a consistently larger size at any given CES value.

Tumour size as predictor of functional activity

Measurement of both urinary or plasma MNs and tumour size were available in 36 patients (19 patients with confirmed PHEO/PGL, 16 patients with suspected PHEO/PGL, 1 patient with confirmed PHEO and suspected contralateral disease recurrence) harbouring a total of 40 tumours. Fifty-three percent ($n = 21$) of the tumours were biochemically active and 47% ($n = 19$) were inactive. Median size of active tumours was significantly larger at 4.46 ± 3.31 cm (range 0.8–11 cm) compared with inactive tumours at 1.17 ± 0.40 cm (range 0.50–2.1 cm) at the time of last follow-up prior to surgery ($P < 0.0001$). Logistic regression analysis revealed a 50% probability for a positive/negative biochemical test result at 1.78 cm tumour diameter (Fig. 3), and ROC curve analysis yielded a similar optimum threshold value of 1.6 cm. The respective SE/SP pair was 76.2% (CI 57.1–90.5%) and 89.5% (CI 73.7–100%), and AUC was 0.865 (CI 0.7454–0.9839) (not shown).

Discussion

It is increasingly being recognised that PHEO/PGL incidentally detected in patients with VHL disease by routine surveillance imaging for serious organ manifestations are often small-sized and biochemically inactive. The clinical significance of these tumours is uncertain due to their generally slow growth tendency

Figure 3

Tumour size and probability of a positive/negative metanephrine (MN) result. The intercept of the line of 50% balanced probability for a positive/negative MNs test result (solid horizontal line) with the sigmoidal regression curve coincides with the inflection point. The corresponding threshold tumour size value (1.78 cm) is indicated on the *x*-axis (solid vertical line). The respective value derived from ROC analysis (1.60 cm) is also shown (dashed vertical line). Boxplots represent the size distribution of tumours with a positive (upper) or negative (lower) biochemical test result. Multiple tumours and recurrent disease were treated as outlined in the 'Materials and methods' section. For numbers of patients with histologically proven and suspected PHEO/PGL see 'Results' section.

and low risk of malignant transformation. This insight has not yet, however, found adequate reflection in existing surveillance recommendations. Here we provide evidence that recognition of basic qualities of this particular group of tumours could help optimise follow-up and treatment decisions in afflicted patients.

In this study, we first show that common features of PHEO/PGL in a large series of VHL patients are in general agreement with existing experience. Thus, similar to observations by others the overall incidence of PHEO/PGL was close to 20%. Males were more frequently affected than females, and mean age at diagnosis was 32 ± 16 years. Furthermore, most PHEO/PGL were localised to the adrenals, 46% of the tumours were smaller than 1.5 cm in diameter at initial diagnosis, tumours were often multiple and recurrent, and malignant (metastatic) disease was rare accounting for less than 1.0% of cases (3, 5, 6, 7, 25).

Of particular mention is, however, the fact that progressive disease consistent with RECIST was found in only 7 of 21 patients with small-sized (≤1.5 cm) PHEO/PGL where accurate measurements at diagnosis and prior to surgery were available. In these patients, tumour size had increased by 76% from 1.47 ± 0.85 to 2.59 ± 1.03 cm within 2.5 years on average when surgery was carried out. Yet with this exception, progressive PHEO/PGL was indistinguishable from those of patients who had stable disease applying current diagnostic standards, including

mutation analyses (19, 30). Obviously, therefore, it is not presently possible to make any meaningful predictions about the future course of the disease in the individual patient with a newly detected VHL-related small (≤1.5 cm) incidentaloma suspicious of PHEO/PGL, thereby emphasising the importance of regular imaging. On the other hand, our results just as clearly show that depending on initial tumour size and emerging growth dynamics at one year after diagnosis it may be reasonable and safe to prolong imaging intervals according to the actual needs of a substantial number of these patients.

Closely linked with the issue of imaging is the question of how to proceed with VHL patients harbouring a non-functioning incidentaloma. In our study 41% (n=54) of tumours in 77 patients with proven PHEO/PGL (1.43 tumours/patient) and 48 patients with suspected PHEO/PGL (1.29 tumours/patient) were functionally inactive. These 'biochemically silent' tumours create a diagnostic dilemma in several respects. First, confirmatory functional imaging is unreliable when (urinary) catecholamines are not elevated (6). Second, because there is a high incidence of PHEO/PGL in VHL patients with incidentaloma even repeatedly negative test results will not allow to rule out the disease with acceptable certainty (6, 37, 38). Finally, chances for a positive biochemical result upon periodic retesting are small given the fact that a direct relation exists between tumour size and secretory activity (6, 39) and that most of these tumours have a slow growth tendency (6, 31) which was also observed in the present study. Thus, in order to keep the proportion of unnecessary tests as low as possible, it would certainly be helpful to have a means of defining more closely those patients where the probability of a negative respectively positive biochemical result predominates. Our findings show that tumour size is a useful proxy in this regard. Modelling with logistic regression analysis revealed a balanced probability of 50% for a negative/positive MN test result at a tumour diameter of 1.80 cm and ROC curve analysis yielded a similar threshold value of 1.60 cm. In the light of these results, it is important to recall that in our study 46% of newly detected incidentalomas measured less than 1.5 cm in diameter at initial diagnosis. From this, it seems possible to reduce the frequency of testing in greater than one-half of the patients harbouring such a small-sized tumour since results will be negative both at initial diagnosis and in all subsequent tests except for a defined subgroup with progressive disease (see previously), in whom test results would be expected to eventually turn positive.

A third point of critical importance concerns the diagnostic value of disease-related signs and symptoms in

VHL patients with radiological suspicion of PHEO/PGL. In general, the clinical presentation of PHEO/PGL is highly variable, often nonspecific, and the predictive properties of the individual signs and symptoms for the presence of these tumours are mostly insufficient (40, 41). For instance, the probability of PHEO/PGL is less than 0.5% in the hypertensive population (37, 42) and there is only a 1.0% chance the patient has the tumour in the presence of the characteristic but rare triad headaches, palpitations and profuse sweating (43, 44). On the other hand, the occurrence of the symptom triad in patients diagnosed with hypertension increases the probability of the disease to nearly 6% (41). Therefore, various attempts have been made in the past at combining selected history data and pertinent physical findings common in patients with PHEO/PGL in order to enhance the diagnostic significance of this clinical information (45, 46, 47). Compared to sporadic PHEO/PGL, however, there is an about 20-fold higher probability (<5% vs ~96%) of PHEO/PGL in VHL patients with a newly detected incidentaloma (48). Under these conditions, additional clinical information will not contribute to establishing the diagnosis. Instead, the coincidental occurrence of signs or symptoms suggestive of PHEO/PGL might even be perceived as disease manifestation and prompt inappropriate therapeutic decisions particularly in patients with small-sized tumours where confirmatory biochemical evidence often cannot be obtained (see previously). In fact, in our series of VHL patients operated on because of PHEO/PGL a different tumour entity was diagnosed by histopathology in 4.5% ($n = 7$) of cases indicating that the risk of diagnostic error is neither merely hypothetical nor negligible. The scoring system presented in the current study was developed to help minimise this risk by providing an objective scale for the presence or absence of signs and symptoms of disease. The score is easy to perform in clinical practice and its discriminative properties appear high. Moreover, internal consistency of the score was corroborated by the results of multiple linear regression analysis which showed a strong positive correlation between score value and tumour size both with and without consideration of secretory activity. Hence, the score seems to adequately fulfil major criteria necessary for diagnostic tests. Obviously, however, further prospective studies are needed with the score before finally judging its usefulness and validity.

One last aspect deserves attention. If follow-up testing revealed slightly to moderately elevated (<four-fold above the upper reference limit) normetanephrine (NMN) plasma concentrations (49) in a VHL patient with a small-sized non-functioning incidentaloma or if this test result

was obtained at initial diagnosis, should this patient be scheduled for surgery? There is no simple answer to this question despite a probability of PHEO/PGL verging on certainty (positive predictive value 0.998; (50)), since VHL patients are at a substantially increased risk of loss of both adrenal glands and subsequent dependency on lifelong hormone substitution therapy owing to the high incidence of bilateral, multiple, and recurrent PHEO/PGL associated with this tumour syndrome compared with other hereditary and sporadic PHEO/PGL (see also Table 2). This risk could immediately materialise with the decision for surgery and must be weighted against the indeterminate (32) but likely much smaller risk of a sudden catecholaminergic crisis if surgery is postponed in such a patient with otherwise no clinical symptoms compatible with PHEO/PGL and a stable tumour size. Conversely, the presence or recent onset of clinical signs and symptoms of PHEO/PGL and/or radiological evidence of tumour growth would considerably add to the odds of unforeseeable complications of excessive catecholamine secretion in this same patient, and in those circumstances, timely resection of the tumour should be pursued. In our study mean tumour size was significantly smaller in the group of patients with suspected compared to proven PHEO/PGL (sum diameter 1.49 ± 1.36 cm vs 3.4 ± 2.7 cm), the overwhelming proportion of patients was asymptomatic (94% vs 65%) and in 42% of patients with suspected PHEO/PGL where biochemical results were available the test was positive. From this, it seems that a composite tumour size within the limits specified above at biochemical diagnosis of PHEO/PGL and absence of clinical symptoms would indeed signal a low risk of PHEO/PGL related complications and justify a watchful waiting attitude unless emerging clinical and imaging evidence indicates progressive disease. Hence, an individualised treatment strategy might also be feasible in patients with small-sized incidentalomas and biochemical evidence of PHEO/PGL upon follow-up testing or at initial diagnosis and help make appropriate therapy decisions with the aim to avoid unnecessary surgery and preserve endogenous adrenal gland function.

Our study has limitations inherent to the retrospective observational study design. Thus, clinical signs and symptoms were not uniformly recorded for all patients and information on imaging and biochemical tests was also incomplete most likely reflecting the great challenges encountered in the care of VHL patients where the multiplicity of regular diagnostic tests for the prevention of potentially life-threatening sequelae of serious disease manifestations may at times lead to a more liberal attitude

towards others. In order to mitigate the effect of missing data, we adopted the strategy by first fitting all regression models by backward selection of a maximum model of complete cases with regard to pertinent baseline variables and then sequentially adding individual parameters of interest from the relevant available cases as described in the 'Materials and methods' section. With this approach, we found a high degree of consistency between results. Therefore we are confident in the informative value of our analyses. We cannot, however, entirely rule out analytical error due to small sample size. This holds particularly true for the results obtained on tumour growth dynamics which relied on only 21 patients with small-sized incidentalomas. The possibility of underrating the proportion of patients with progressive disease must, therefore, be borne in mind. Yet, in view of the large total number of patients with small-sized incidentalomas in the whole study cohort, there would still remain a substantial number of patients with stable disease in absolute terms who could ultimately benefit from longer time intervals between CT/MR imaging and a reduced frequency of biochemical testing as suggested in our study.

In conclusion results of the present study support the case for high-resolution CT/MR measurement of tumour size and measurement of MNs in urine or plasma in VHL patients with newly detected non-functioning small-sized incidentalomas at initial diagnosis, at one year following diagnosis and at individualised intervals thereafter depending on evolving growth dynamics and secretory activity. Also, physical examination should include standardised assessment of disease-related signs and symptoms at all visits. It is anticipated that implementation of these surveillance principles will help reduce the risks of diagnostic error and surgery and the burden of tests and associated distress in many of these particular patients.

References

1 Haase VH. The VHL tumor suppressor: master regulator of HIF. *Current Pharmaceutical Design* 2009 **15** 3895–3903. (https://doi.org/10.2174/138161209789649394)

2 Neumann HP & Wiestler OD. Clustering of features of von Hippel-Lindau syndrome: evidence for a complex genetic locus. *Lancet* 1991 **337** 1052–1054. (https://doi.org/10.1016/0140-6736(91)91705-y)

3 Aufforth RD, Ramakant P, Sadowski SM, Mehta A, Trebska-McGowan K, Nilubol N, Pacak K & Kebebew E. Pheochromocytoma screening initiation and frequency in von Hippel-Lindau syndrome. *Journal of Clinical Endocrinology and Metabolism* 2015 **100** 4498–4504. (https://doi.org/10.1210/jc.2015-3045)

4 Neumann HP & Bender BU. Genotype-phenotype correlations in von Hippel-Lindau disease. *Journal of Internal Medicine* 1998 **243** 541–545. (https://doi.org/10.1046/j.1365-2796.1998.00336.x)

5 Ong KR, Woodward ER, Killick P, Lim C, Macdonald F & Maher ER. Genotype-phenotype correlations in von Hippel-Lindau disease. *Human Mutation* 2007 **28** 143–149. (https://doi.org/10.1002/humu.20385)

6 Walther MM, Reiter R, Keiser HR, Choyke PL, Venzon D, Hurley K, Gnarra JR, Reynolds JC, Glenn GM, Zbar B, *et al*. Clinical and genetic characterization of pheochromocytoma in von Hippel-Lindau families: comparison with sporadic pheochromocytoma gives insight into natural history of pheochromocytoma. *Journal of Urology* 1999 **162** 659–664. (https://doi.org/10.1097/00005392-199909010-00004)

7 Neumann HP, Bausch B, McWhinney SR, Bender BU, Gimm O, Franke G, Schipper J, Klisch J, Altehoefer C, Zerres K, *et al*. Germ-line mutations in nonsyndromic pheochromocytoma. *New England Journal of Medicine* 2002 **346** 1459–1466. (https://doi.org/10.1056/NEJMoa020152)

8 Dahia PL, Ross KN, Wright ME, Hayashida CY, Santagata S, Barontini M, Kung AL, Sanso G, Powers JF, Tischler AS, *et al*. A HIF1alpha regulatory loop links hypoxia and mitochondrial signals in pheochromocytomas. *PLoS Genetics* 2005 **1** 72–80. (https://doi.org/10.1371/journal.pgen.0010008)

9 Eisenhofer G, Huynh TT, Pacak K, Brouwers FM, Walther MM, Linehan WM, Munson PJ, Mannelli M, Goldstein DS & Elkahloun AG. Distinct gene expression profiles in norepinephrine- and epinephrine-producing hereditary and sporadic pheochromocytomas: activation of hypoxia-driven angiogenic pathways in von Hippel-Lindau syndrome. *Endocrine-Related Cancer* 2004 **11** 897–911. (https://doi.org/10.1677/erc.1.00838)

10 Favier J, Briere JJ, Burnichon N, Riviere J, Vescovo L, Benit P, Giscos-Douriez I, De Reynies A, Bertherat J, Badoual C, *et al*. The Warburg effect is genetically determined in inherited pheochromocytomas. *PLoS ONE* 2009 **4** e7094. (https://doi.org/10.1371/journal.pone.0007094)

11 Eisenhofer G, Walther MM, Huynh TT, Li ST, Bornstein SR, Vortmeyer A, Mannelli M, Goldstein DS, Linehan WM, Lenders JW, *et al*. Pheochromocytomas in von Hippel-Lindau syndrome and multiple endocrine neoplasia type 2 display distinct biochemical and clinical phenotypes. *Journal of Clinical Endocrinology and Metabolism* 2001 **86** 1999–2008. (https://doi.org/10.1210/jcem.86.5.7496)

12 Burnichon N, Vescovo L, Amar L, Libe R, de Reynies A, Venisse A, Jouanno E, Laurendeau I, Parfait B, Bertherat J, *et al*. Integrative genomic analysis reveals somatic mutations in pheochromocytoma and paraganglioma. *Human Molecular Genetics* 2011 **20** 3974–3985. (https://doi.org/10.1093/hmg/ddr324)

13 Galan SR & Kann PH. Genetics and molecular pathogenesis of pheochromocytoma and paraganglioma. *Clinical Endocrinology* 2013 **78** 165–175. (https://doi.org/10.1111/cen.12071)

14 Lopez-Jimenez E, Gomez-Lopez G, Leandro-Garcia LJ, Munoz I, Schiavi F, Montero-Conde C, de Cubas AA, Ramires R, Landa I, Leskela S, *et al*. Research resource: transcriptional profiling reveals different pseudohypoxic signatures in SDHB and VHL-related pheochromocytomas. *Molecular Endocrinology* 2010 **24** 2382–2391. (https://doi.org/10.1210/me.2010-0256)

15 Vicha A, Musil Z & Pacak K. Genetics of pheochromocytoma and paraganglioma syndromes: new advances and future

treatment options. *Current Opinion in Endocrinology, Diabetes, and Obesity* 2013 **20** 186–191. (https://doi.org/10.1097/MED.0b013e32835fcc45)

16 Benn DE, Gimenez-Roqueplo AP, Reilly JR, Bertherat J, Burgess J, Byth K, Croxson M, Dahia PL, Elston M, Gimm O, *et al*. Clinical presentation and penetrance of pheochromocytoma/paraganglioma syndromes. *Journal of Clinical Endocrinology and Metabolism* 2006 **91** 827–836. (https://doi.org/10.1210/jc.2005-1862)

17 Clark GR, Sciacovelli M, Gaude E, Walsh DM, Kirby G, Simpson MA, Trembath RC, Berg JN, Woodward ER, Kinning E, *et al*. Germline FH mutations presenting with pheochromocytoma. *Journal of Clinical Endocrinology and Metabolism* 2014 **99** E2046–E2050. (https://doi.org/10.1210/jc.2014-1659)

18 Därr R, Nambuba J, Del Rivero J, Janssen I, Merino M, Todorovic M, Balint B, Jochmanova I, Prchal JT, Lechan RM, *et al*. Novel insights into the polycythemia-paraganglioma-somatostatinoma syndrome. *Endocrine-Related Cancer* 2016 **23** 899–908. (https://doi.org/10.1530/ERC-16-0231)

19 Maher ER, Neumann HP & Richard S. Von Hippel-Lindau disease: a clinical and scientific review. *European Journal of Human Genetics* 2011 **19** 617–623. (https://doi.org/10.1038/ejhg.2010.175)

20 Timmers HJ, Kozupa A, Eisenhofer G, Raygada M, Adams KT, Solis D, Lenders JW & Pacak K. Clinical presentations, biochemical phenotypes, and genotype-phenotype correlations in patients with succinate dehydrogenase subunit B-associated pheochromocytomas and paragangliomas. *Journal of Clinical Endocrinology and Metabolism* 2007 **92** 779–786. (https://doi.org/10.1210/jc.2006-2315)

21 Wilding A, Ingham SL, Lalloo F, Clancy T, Huson SM, Moran A & Evans DG. Life expectancy in hereditary cancer predisposing diseases: an observational study. *Journal of Medical Genetics* 2012 **49** 264–269. (https://doi.org/10.1136/jmedgenet-2011-100562)

22 Al-Hawary MM, Francis IR & Korobkin M. Non-invasive evaluation of the incidentally detected indeterminate adrenal mass. *Best Practice and Research: Clinical Endocrinology and Metabolism* 2005 **19** 277–292. (https://doi.org/10.1016/j.beem.2004.09.002)

23 Kaji P, Carrasquillo JA, Linehan WM, Chen CC, Eisenhofer G, Pinto PA, Lai EW & Pacak K. The role of 6-[18F]fluorodopamine positron emission tomography in the localization of adrenal pheochromocytoma associated with von Hippel-Lindau syndrome. *European Journal of Endocrinology* 2007 **156** 483–487. (https://doi.org/10.1530/EJE-06-0712)

24 Young Jr WF. Clinical practice. The incidentally discovered adrenal mass. *New England Journal of Medicine* 2007 **356** 601–610. (https://doi.org/10.1056/NEJMcp065470)

25 Yao M, Shinohara N, Yamazaki I, Tamura K & Shuin T. Von Hippel-Lindau disease-associated pheochromocytoma: epidemiology, clinical characteristics, and screening and surveillance protocols in Japan. *Journal of Translational Medicine and Epidemiology* 2014 **2** 1014.

26 Binderup ML, Bisgaard ML, Harbud V, Moller HU, Gimsing S, Friis-Hansen L, Hansen Tv, Bagi P, Knigge U, Kosteljanetz M, *et al*. Von Hippel-Lindau disease (vHL). National clinical guideline for diagnosis and surveillance in Denmark. 3rd edition. *Danish Medical Journal* 2013 **60** B4763.

27 Hes FJ, van der Luijt RB & Lips CJ. Clinical management of von Hippel-Lindau (VHL) disease. *Netherlands Journal of Medicine* 2001 **59** 225–234. (https://doi.org/10.1016/s0300-2977(01)00165-6)

28 Kruizinga RC, Sluiter WJ, de Vries EG, Zonnenberg BA, Lips CJ, van der Horst-Schrivers AN, Walenkamp AM & Links TP. Calculating optimal surveillance for detection of von Hippel-Lindau-related manifestations. *Endocrine-Related Cancer* 2014 **21** 63–71. (https://doi.org/10.1530/ERC-13-0308)

29 Rednam SP, Erez A, Druker H, Janeway KA, Kamihara J, Kohlmann WK, Nathanson KL, States LJ, Tomlinson GE, Villani A, *et al*. Von Hippel-Lindau and hereditary pheochromocytoma/paraganglioma syndromes: clinical features, genetics, and surveillance recommendations in childhood. *Clinical Cancer*

Research 2017 **23** e68–e75. (https://doi.org/10.1158/1078-0432.CCR-17-0547)

30 VHL Alliance. VHLA suggested active surveillance guidelines (updated 4/24/2020). Boston, MA, USA: VHL Alliance, 2020. (available at: https://www.vhl.org/wp-content/uploads/2020/07/Active-Surveillance-Guidelines-2020-2.pdf)

31 Lenders JW, Eisenhofer G, Mannelli M & Pacak K. Phaeochromocytoma. *Lancet* 2005 **366** 665–675. (https://doi.org/10.1016/S0140-6736(05)67139-5)

32 Mannelli M, Lenders JW, Pacak K, Parenti G & Eisenhofer G. Subclinical phaeochromocytoma. *Best Practice and Research: Clinical Endocrinology and Metabolism* 2012 **26** 507–515. (https://doi.org/10.1016/j.beem.2011.10.008)

33 Mundschenk J, Dieterich KD, Kopf D, Höppner W & Lehnert H. Phäochromozytom: Klinik, Diagnostik und Therapie. *Deutsches Ärzteblatt International* 2001 **98** 2502–2510.

34 Cappelleri JC, Jason Lundy J & Hays RD. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clinical Therapeutics* 2014 **36** 648–662. (https://doi.org/10.1016/j.clinthera.2014.04.006)

35 Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, *et al*. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *European Journal of Cancer* 2009 **45** 228–247. (https://doi.org/10.1016/j.ejca.2008.10.026)

36 DeLong ER, DeLong DM & Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988 **44** 837–845. (https://doi.org/10.2307/2531595)

37 Anderson Jr GH, Blakeman N & Streeten DH. The effect of age on prevalence of secondary forms of hypertension in 4429 consecutively referred patients. *Journal of Hypertension* 1994 **12** 609–615. (https://doi.org/10.1097/00004872-199405000-00015)

38 Aprill BS, Drake 3rd AJ, Lasseter DH & Shakir KM. Silent adrenal nodules in von Hippel-Lindau disease suggest pheochromocytoma. *Annals of Internal Medicine* 1994 **120** 485–487. (https://doi.org/10.7326/0003-4819-120-6-199403150-00006)

39 Eisenhofer G, Lenders JW, Goldstein DS, Mannelli M, Csako G, Walther MM, Brouwers FM & Pacak K. Pheochromocytoma catecholamine phenotypes and prediction of tumor size and location by use of plasma free metanephrines. *Clinical Chemistry* 2005 **51** 735–744. (https://doi.org/10.1373/clinchem.2004.045484)

40 Manger WM. The protean manifestations of pheochromocytoma. *Hormone and Metabolic Research* 2009 **41** 658–663. (https://doi.org/10.1055/s-0028-1128139)

41 Plouin PF, Degoulet P, Tugaye A, Ducrocq MB & Ménard J. A semiological study of 2585 patients, including. *Nouvelle Presse Medicale* 1981 **10** 869–872.

42 Ariton M, Juan CS & AvRuskin TW. Pheochromocytoma: clinical observations from a Brooklyn tertiary hospital. *Endocrine Practice* 2000 **6** 249–252. (https://doi.org/10.4158/EP.6.3.249)

43 Falhammar H, Kjellman M & Calissendorff J. Initial clinical presentation and spectrum of pheochromocytoma: a study of 94 cases from a single center. *Endocrine Connections* 2018 **7** 186–192. (https://doi.org/10.1530/EC-17-0321)

44 Kopetschke R, Slisko M, Kilisli A, Tuschy U, Wallaschofski H, Fassnacht M, Ventz M, Beuschlein F, Reincke M, Reisch N, *et al*. Frequent incidental discovery of phaeochromocytoma: data from a German cohort of 201 phaeochromocytoma. *European Journal of Endocrinology* 2009 **161** 355–361. (https://doi.org/10.1530/EJE-09-0384)

45 Black HR & Bursten SL. A clinical scoring system for detection of patients with pheochromocytomas. *Yale Journal of Biology and Medicine* 1984 **57** 259–272.

46 Soltani A, Pourian M & Davani BM. Does this patient have pheochromocytoma? A systematic review of clinical signs and

symptoms. *Journal of Diabetes and Metabolic Disorders* 2015 **15** 6. (https://doi.org/10.1186/s40200-016-0226-x)

47 Stein PP & Black HR. A simplified diagnostic approach to pheochromocytoma. A review of the literature and report of one institution's experience. *Medicine* 1991 **70** 46–66. (https://doi.org/10.1097/00005792-199101000-00004)

48 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. *Journal of Clinical Endocrinology and Metabolism* 2000 **85** 637–644. (https://doi.org/10.1210/jcem.85.2.6372)

49 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. *Journal of Clinical Endocrinology and Metabolism* 2003 **88** 2656–2666. (https://doi.org/10.1210/jc.2002-030005)

50 Därr R, Kuhn M, Bode C, Bornstein SR, Pacak K, Lenders JWM & 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. (https://doi.org/10.1007/s12020-017-1300-y)