Urine steroid metabolomics for the differential diagnosis of adrenal incidentalomas in the EURINE-ACT study
ENSAT EURINE-ACT Investigators

DOI: 10.1016/S2213-8587(20)30218-7

License: Creative Commons: Attribution (CC BY)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 22. Aug. 2021
Urine steroid metabolomics for the differential diagnosis of adrenal incidentalomas in the EURINE-ACT study: a prospective test validation study


Summary

Background Cross-sectional imaging regularly results in incidental discovery of adrenal tumours, requiring exclusion of adrenocortical carcinoma (ACC). However, differentiation is hampered by poor specificity of imaging characteristics. We aimed to validate a urine steroid metabolomics approach, using steroid profiling as the diagnostic basis for ACC.

Methods We did a prospective multicentre study in adult participants (age ≥18 years) with newly diagnosed adrenal masses. We assessed the accuracy of diagnostic imaging strategies based on maximum tumour diameter (≥4 cm vs <4 cm), imaging characteristics (positive vs negative), and urine steroid metabolomics (low, medium, or high risk of ACC), separately and in combination, using a reference standard of histopathology and follow-up investigations. With respect to imaging characteristics, we also assessed the diagnostic utility of increasing the unenhanced CT tumour attenuation threshold from the recommended 10 HU to 20 HU.

Findings Of 2169 participants recruited between Jan 17, 2011, and July 15, 2016, we included 2017 from 14 specialist centres in 11 countries in the final analysis. 98 (4·9%) had histopathologically or clinically and biochemically confirmed ACC. Tumours with diameters of 4 cm or larger were identified in 488 participants (24·2%), including 96 of the 98 with ACC (positive predictive value [PPV] 19·7%, 95% CI 16·2–23·5). For imaging characteristics, increasing the unenhanced CT tumour attenuation threshold to 20 HU from the recommended 10 HU increased specificity for ACC (80·0% [95% CI 77·9–82·0] vs 64·0% [61·4–66·4]) while maintaining sensitivity (99·0% [94·4–100·0] vs 100·0% [96·3–100·0]; PPV 19·7%, 16·3–23·5). A urine steroid metabolomics result indicated high risk of ACC had a PPV of 34·6% (95% CI 28·6–41·0). When the three tests were combined, in the order of tumour specificity for ACC (80·0% [95% CI 77·9–82·0]), increasing the unenhanced CT tumour attenuation threshold to 20 HU increased the unenhanced CT tumour attenuation threshold from the recommended 10 HU to 20 HU.

Interpretation An unenhanced CT tumour attenuation cutoff of 20 HU should replace that of 10 HU for exclusion of ACC. A triple test strategy of tumour diameter, imaging characteristics, and urine steroid metabolomics improves detection of ACC, which could shorten time to surgery for patients with ACC and help to avoid unnecessary surgery in patients with benign tumours.

Funding European Commission, UK Medical Research Council, Wellcome Trust, and UK National Institute for Health Research, US National Institutes of Health, the Claire Khan Trust Fund at University Hospitals Birmingham Charities, and the Mayo Clinic Foundation for Medical Education and Research.

Copyright © 2020 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Introduction Adrenal masses are discovered incidentally in about 5% of cross-sectional imaging examinations.1,2 The prevalence of these so-called incidentalomas increases with age and is estimated to be about 3% among people aged 40 years and 10% among those aged 70 years.3 Because of widespread use of CT and MRI, the number of adrenal incidentalomas identified is increasing...
The discovery of an adrenal mass prompts two major questions to be addressed by the diagnostic investigations. The first question is whether the tumour causes adrenal hormone excess, which requires exclusion of pheochromocytoma, Cushing’s syndrome, and primary aldosteronism via biochemical testing. The second question is whether the adrenal mass is an adenocortical carcinoma (ACC). These tumours account for 2–12% of adrenal incidentalomas and are mainly found in the elderly. Detection and exclusion of ACC is a challenging health-care issue.

Research in context

Evidence before this study

The discovery of an adrenal mass requires further assessment to determine whether the tumour is a malignant adrenocortical carcinoma (ACC) or a benign adrenocortical adenoma (ACA). Imaging tests are the mainstay in detection of ACC, but the available methods have poor diagnostic accuracy. In 2016 we did a systematic review and meta-analysis of the accuracy of imaging modalities. We searched MEDLINE, Embase, the Cochrane Central Register of Controlled Trials, the Science Citation Index, the Conference Proceedings Citation Index, and Zetoc for reports published from Jan 1, 1990, to Aug 31, 2015, to identify all studies that had assessed the diagnostic accuracy of imaging methods employed in the differential diagnosis of adrenal masses (CT, MRI, and ¹⁸F-fluorodeoxyglucose [¹⁸F-FDG] PET) in comparison with histology or imaging-based follow-up as the reference standard. The search terms used for these database searches were reported in that study. Based on data from only two eligible studies involving 102 participants, an unenhanced CT tumour attenuation value of more than 10 Hounsfield units (HU) was shown to have high sensitivity (100%) but poor specificity (72%) for the detection of ACC. Studies assessing ¹⁸F-FDG PET, MRI with chemical shift analysis, and CT with contrast washout had even smaller sample sizes (n=25–75) and generated sensitivity and specificity estimates with very wide CIs. In 2019 we updated our previous search with a PubMed search using the same search terms and covering the dates Sept 1, 2015–Jan 31, 2019. We found no studies that altered the conclusions of the 2016 systematic review and meta-analysis, nor any newly published meta-analyses addressing the same question. We did a retrospective proof-of-principle study in 45 participants with ACC and 102 with ACA, in which we showed that urine steroid metabolomics (ie, mass spectrometry-based steroid metabolite profiling of 24 h urine samples combined with machine-learning-based data analysis) could detect ACC with high sensitivity and specificity. Although similar diagnostic accuracies were reported in several other retrospective studies, we did not identify any previous prospective test validation studies.

Added value of this study

In this prospective, multicentre, test validation study, we assessed the diagnostic accuracy of urine steroid metabolomics for the detection of ACC in patients with newly diagnosed adrenal masses compared with routinely employed imaging strategies (measurement of tumour size and assessment of imaging characteristics). As a single test, urine steroid metabolomics had a higher positive predictive value (PPV) than either imaging test (34.6% [95% CI 28.6–41.0] vs 19.7% [16.2–23.5] for tumour size and 19.7% [16.3–23.5] for imaging characteristics). Additionally, in a comparison with the standard cutoff of 10 HU for unenhanced CT tumour attenuation, a cutoff of 20 HU increased specificity from 64% to 80% while sensitivity remained similar. The best diagnostic performance was provided by the combination of tumour diameter greater than 4 cm, unenhanced CT tumour attenuation greater than 20 HU, and urine steroid metabolomics indicating a high risk of ACC, which provided a PPV of 76.4% and a negative predictive value of 99.7% for ACC.

Implications of all the available evidence

This study prospectively validated the diagnostic accuracy of urine steroid metabolomics seen in retrospective studies. Use of radiation-free, non-invasive urine steroid metabolomics has a higher PPV than two standard imaging tests, and best performance was seen with the combination of all three tests. Implementation of urine steroid metabolomics in the routine diagnostic assessment of newly discovered adrenal masses could reduce the number of imaging procedures required to diagnose ACC and avoid unnecessary surgery of benign adrenal tumours, potentially yielding beneficial effects with respect to patient burden and health-care costs.

worldwide. In the USA, more than 100 million CT and MRI examinations are done per year, which could potentially reveal around 5 million adrenal masses. Diagnostic assessment and therapeutic management of incidentalomas is a challenging health-care issue.

The discovery of an adrenal mass prompts two major questions to be addressed by the diagnostic investigations. The first question is whether the tumour causes adrenal hormone excess, which requires exclusion of pheochromocytoma, Cushing’s syndrome, and primary aldosteronism via biochemical testing. The second question is whether the adrenal mass is an adenocortical carcinoma (ACC). These tumours account for 2–12% of adrenal incidentalomas. Prognosis of ACC is poor and cure is achievable only with early detection and surgery. Multimodal and repeated imaging is routinely used to detect and exclude ACC, despite an insufficient evidence base to support this approach due to small studies in heterogeneous populations and with poor reference standards. Based on available data, an unenhanced CT tumour attenuation of less than 10 Hounsfield units (HU) is thought to be indicative of a lipid-rich mass, which is a common characteristic of benign adrenal tumours. European guidelines, therefore, recommend that unenhanced CT should be the primary imaging test to exclude ACC. However, the 10 HU cutoff has poor specificity for malignancy because of the existence of many lipid-poor benign adrenal masses. Patients therefore frequently undergo repeated imaging by multiple modalities and the unnecessary surgical removal of adrenal masses that are revealed postoperatively to be benign. Urine steroid metabolomics is the combination of mass spectrometry-based urinary steroid metabolite
profiling and machine-learning-based data analysis. In a retrospective proof-of-principle study, we showed that this approach was highly promising for the detection of ACC, with high sensitivity (90%) and specificity (88%). Similar findings were reported in several subsequent retrospective studies.

To validate the use of urine steroid metabolomic testing in this context, we did the prospective, multicentre Evaluation of Urine Steroid Metabolomics in the Differential Diagnosis of Adrenocortical Tumours (EURINE-ACT) study in adult participants with newly diagnosed adrenal masses. We investigated the diagnostic accuracy of urine steroid metabolomics alone and in combination with standard imaging protocols, and compared our findings with histopathology and clinical and imaging follow-up investigations as the reference standard.

Methods
Study design and participants
EURINE-ACT was a prospective test validation study performed according to the STARD guidelines for studies of diagnostic accuracy (appendix pp 36–37) and done in adult participants (age ≥18 years) with a newly identified adrenal mass of more than 1 cm diameter. Exclusion criteria were biochemical evidence of pheochromocytoma (appendix p 5), pregnancy, lactation, and current or recent (<6 months) intake of drugs known to alter steroid synthesis or metabolism. Patients with an adrenal mass discovered during imaging for cancer staging or monitoring were also not eligible.

Participants were recruited through specialist centres participating in the European Network for the Study of Adrenal Tumours (ENSAT); appendix p 4. The study was advertised to all ENSAT members, and 21 centres in 14 countries (Brazil, Croatia, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Poland, Portugal, Serbia, the UK, and the USA) agreed to participate and initiated enrolment. We asked centres to recruit prospectively consecutive eligible individuals willing to participate (ie, non-selective recruitment), from Jan 17, 2011, to July 15, 2016.

All participating centres obtained local ethics approval for recording of pseudonymised and standardised data in the ENSAT registry relevant to the study (demographic characteristics, method of tumour identification, tumour diameter and imaging characteristics, endocrine testing results, clinical and radiological follow-up data, surgery details, and histopathology data) and for collection and use of participant-related biomaterial (appendix pp 4–5). All participants provided written informed consent before inclusion.

Reference standard
The reference standard for ACC was based on histopathology or, alternatively, the presence of a large adrenal tumour and mixed steroid excess typical of ACC, with no other feasible alternative diagnosis. Adrenal masses were classified as benign based on histopathology after surgical removal, or, in those not removed, by lack of growth on imaging after 6 months or clinical follow-up of at least 12 months (appendix p 23).

Imaging assessments
Diagnostic investigation for adrenal tumours by imaging was done as part of routine care at the clinical centres in accordance with their standard guidelines. Local centres made decisions about the need for adrenalectomy (and in rare cases adrenal biopsy) based on imaging, without any access to urine steroid metabolomics data. The imaging index tests we recorded were maximum tumour diameter at the time of discovery and imaging characteristics. Tumour diameter of 4 cm or greater indicated suspicion of ACC.

Imaging characteristics regarded as suspicious of ACC were as follows (in order of ranking): unenhanced CT tumour tissue attenuation (<10 HU, 10–20 HU, or >20 HU in homogeneous tumours or heterogeneous tumours precluding HU measurement); MRI chemical shift analysis with no loss of signal intensity in the tumour area on out-of-phase images; ¹⁸F-fluorodeoxyglucose (¹⁸F-FDG) PET with a tumour standardised uptake value higher than that in the liver; follow-up imaging showing an increase of the maximum diameter of at least 20% 6 months or more after the index scan; and CT contrast washout assessment showing absolute contrast washout from less than 60% of the tumour area. If participants underwent multiple imaging modalities, a positive or negative result for ACC was based on the highest ranking imaging characteristic reported.

Urine steroid metabolomics testing
Enrolled participants collected a 24 h urine sample that was used for multistroid profiling by liquid chromatography–tandem mass spectrometry (LC–MS/MS), with quantification of 15 urinary steroid metabolites (appendix p 19) and application of a machine-learning algorithm.

The algorithm was developed by applying generalised matrix learning vector quantisation to steroid excretion data from a prospective cohort of 139 patients with adrenal masses (40 ACC and 99 adrenocortical adenoma [ACA]) measured retrospectively by use of the LC–MS/MS method used in this study (appendix pp 7–9). Based on the distances of the entirety of the steroid metabolome of ACC and ACA prototypes, the generalised matrix learning vector quantisation classifier provides a test outcome score. The corresponding thresholds were selected to ensure the post-test probability of ACC was greater than 65% in the high-risk group and less than 10% in the low-risk group, giving a moderate risk range of 10–65% (appendix p 9).

Urine steroid metabolomics analyses were done after all routine tests in the study centres were completed, but
without access to the reference standard findings or other diagnostic information. This timing also prevented urine steroid metabolomics results being communicated to the study centres and affecting the normal diagnostic process.

Statistical analysis

We aimed to include 2000 participants and expected an ACC rate of 5%, based on the results of our proof-of-principle study (appendix p 4), and loss to follow-up of 10%. We calculated that observing 100 ACC cases would provide 95% sensitivity with a 95% CI range of less than 10% and more than 99% power to detect a difference of 3% in specificity (87% vs 90%) between standard imaging protocols and urine steroid metabolomics at the 5% significance level.

Characteristics of participants are reported for each type of tumour assessed (ACC, other malignant tumours, ACA, and other benign tumours), with data presented as median (IQR) for categorical results or number (%) for continuous data. For each diagnostic test (maximum tumour diameter ≥4 cm vs <4), imaging characteristics [positive vs negative], or urine steroid metabolomics [low, medium, or high risk of ACC]), we computed the percentage of ACC cases with each test result (giving sensitivity for a positive result in the binary tests); the percentage of non-ACC cases with each test result (giving specificity for a negative result in the binary tests); and the likelihood ratio for each test result. Additionally, we assessed the tests in combinations of two (tumour diameter plus imaging characteristics, tumour diameter plus urine steroid metabolomics, and imaging characteristics plus urine steroid metabolomics) and as a triple test strategy (tumour diameter, followed by imaging characteristics followed by urine steroid metabolomics). We also calculated the proportions of participants with ACC who had each test result to estimate the probability of ACC (ie, positive predictive value [PPV] for positive results and 1 minus negative predictive value [NPV] for negative test results). All results are reported with 95% CIs, computed by use of the exact binomial method for proportions and the Wald-based methods for likelihood ratios. Finally, we investigated the diagnostic utility of increasing the unenhanced CT tumour attenuation threshold from the recommended 10 HU to 20 HU.

For urine steroid metabolomics, we calculated the area under the receiver operating characteristic curve (AUROC) with 95% CIs. A sensitivity analysis was done that excluded participants with ACC who had mixed or aberrant steroid excess, clinical presentation of large adrenal mass with extra-adrenal metastases, or bilateral macronodular adrenal hyperplasia and isolated cortisol excess identifying presumed benign tumours.

All statistical analyses were done with Stata version 16 and graphs were created in R.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Between Jan 17, 2011, and July 15, 2016, 2169 eligible participants were recruited from the 21 participating centres and provided 24 h urine samples (figure 1, appendix p 14). However, review of the data showed that only 14 of the 21 centres had a high median annual recruitment rate (33 participants per year) and a median centre-specific proportion of ACC of 3·9%. These 14 centres recruited a total of 2068 participants. By contrast, the seven remaining centres jointly recruited only 101 participants (median annual recruitment rate 6·7), but the median proportion of ACC cases was 35% (appendix p 21). Investigators at these seven sites confirmed that they had recruited selectively, favouring large and suspicious masses, compared with non-selective consecutive recruitment at the other 14 centres.

Therefore, we excluded the 101 participants from these seven centres. An additional 51 participants were excluded from the analysis because of sample loss during storage, transport, or processing; therefore, the final analysis cohort consisted of 2017 participants (figure 1). No participants were lost to follow-up.

The median age of included participants was 59 years, 62% were women, and 84% of the adrenal tumours were discovered incidentally (table 1). Diagnosis by the reference standards was ACC in 98 (4·9%), other malignant tumours in 65 (3·2%), ACA in 1767 (87·6%), and other benign tumours in 87 participants (4·3%; table 1, appendix pp 11, 22). Histopathology results were used to provide the diagnosis for 91 (92·9%) of the ACCs, 65 (100·0%) of other malignant tumours, 370 (20·9%) of ACAs, and 59 (67·8%) of other benign tumours. Clinical and radiological follow-up assessments were used to provide the diagnosis for all remaining tumours (appendix p 23).

Figure 1: Study profile
563 (27.9%) of 2017 participants underwent adrenalectomy, including 370 with ACA and 59 with other benign tumours (appendix p 22). 186 of these ACAs (50·3%) were either non-functioning (n=81) or showed only mild autonomous cortisol secretion (n=105). Thus, 245 (43·5%) of 563 surgically managed participants would not have required adrenalectomy. Even if mild autonomous cortisol secretion was used as an indication for adrenalectomy, which is not current clinical practice, 140 (24·9%) of 563 adrenal masses would still have not required adrenalectomy.

2737 imaging tests were done in the 2017 participants (appendix p 18). A tumour diameter of at least 4 cm was seen in most participants with ACC and other malignant tumours, around two-thirds of those with other benign tumours and nearly 77% of those with ACA (table 1, figure 2, appendix p 24). The highest-ranking imaging characteristics results were obtained by unenhanced CT in 1549 (76.8%), MRI chemical shift analysis in 227 (11·3%), ¹⁸F-FDG PET in 43 (2.1%), follow-up CT in 155, and CT contrast washout assessment in six participants (appendix p 18). All 98 participants with ACC had homogeneous tumours with unenhanced CT tumour attenuation of at least 10 HU or heterogeneous tumours precluding HU measurement (table 1). However, among 1328 participants with ACA, unenhanced CT attenuation was greater than the cutoff of 10 HU (ie, false positive) in 21 (15.1%) of 132 participants with ACA, 59 (77.9%) of 75 with other benign tumours, and 1455 (82.3%) of 1775 with other malignant tumours. With attenuation greater than the cutoff of 20 HU, 97 (99%) participants with ACC and other malignant tumours required adrenalectomy.

The accuracy of urine steroid metabolomics was high (AUROC 94·6%, 95% CI 92·2–96·9; appendix p 29). Changing the cutoff from 10 HU to 200 (15·1%) improved specificity across the whole study cohort (figure 2, appendix p 24). The highest-ranking imaging modalities—unenhanced CT, MRI chemical shift analysis, and ¹⁸F-FDG PET—showed that high risk of ACC and ¹⁸F-FDG PET—showed that high risk of ACC

### Table 1: Clinical characteristics and radiological findings of participants included in the study cohort

<table>
<thead>
<tr>
<th>ACC (n=98)</th>
<th>Other malignant tumours (n=65)</th>
<th>ACA (n=1767)</th>
<th>Other benign tumours (n=87)</th>
<th>Total (n=2017)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>37 (37.8%)</td>
<td>38 (58.5%)</td>
<td>663 (37.5%)</td>
<td>775 (38.4%)</td>
</tr>
<tr>
<td>Women</td>
<td>61 (62.2%)</td>
<td>27 (41.5%)</td>
<td>1104 (62.5%)</td>
<td>1242 (61.6%)</td>
</tr>
<tr>
<td><strong>Age at diagnosis (years)</strong></td>
<td>50 (42-61)</td>
<td>64 (53-70)</td>
<td>59 (50-67)</td>
<td>52 (44-62)</td>
</tr>
<tr>
<td><strong>Incidentally discovered</strong></td>
<td>Yes</td>
<td>42 (49.9%)</td>
<td>56 (86.2%)</td>
<td>1512 (85.6%)</td>
</tr>
<tr>
<td>No*</td>
<td>55 (56.1%)</td>
<td>13 (23.8%)</td>
<td>254 (14.4%)</td>
<td>13 (14.9%)</td>
</tr>
<tr>
<td><strong>Location of tumour</strong></td>
<td>Right adrenal</td>
<td>38 (38.8%)</td>
<td>23 (35.4%)</td>
<td>556 (31.5%)</td>
</tr>
<tr>
<td>Left adrenal</td>
<td>60 (61.2%)</td>
<td>35 (53.8%)</td>
<td>843 (47.7%)</td>
<td>35 (40.2%)</td>
</tr>
<tr>
<td>Both adrenals</td>
<td>0</td>
<td>7 (12.8%)</td>
<td>368 (20.8%)</td>
<td>9 (10.3%)</td>
</tr>
<tr>
<td><strong>Maximum tumour diameter (cm)</strong></td>
<td>&lt;2</td>
<td>0</td>
<td>6 (9.2%)</td>
<td>559 (31.6%)</td>
</tr>
<tr>
<td>&gt;2 to &lt;4</td>
<td>2 (2.0%)</td>
<td>13 (20.0%)</td>
<td>912 (51.6%)</td>
<td>26 (29.9%)</td>
</tr>
<tr>
<td>&gt;4</td>
<td>96 (98.0%)</td>
<td>46 (70.8%)</td>
<td>296 (16.8%)</td>
<td>24 (27.6%)</td>
</tr>
<tr>
<td><strong>Unenhanced CT tumour attenuation</strong></td>
<td>&lt;10 HU*</td>
<td>0</td>
<td>0</td>
<td>905 (51.2%)</td>
</tr>
<tr>
<td>10–20 HU†</td>
<td>1 (1.0%)</td>
<td>1 (1.5%)</td>
<td>223 (12.6%)</td>
<td>9 (10.3%)</td>
</tr>
<tr>
<td>&gt;20 HU‡</td>
<td>32 (32.7%)</td>
<td>38 (58.5%)</td>
<td>200 (11.3%)</td>
<td>30 (34.5%)</td>
</tr>
<tr>
<td><strong>Heterogeneous§</strong></td>
<td>65 (66.3%)</td>
<td>21 (32.3%)</td>
<td>0</td>
<td>1 (1.1%)</td>
</tr>
<tr>
<td><strong>Not done</strong></td>
<td>96 (98.0%)</td>
<td>46 (70.8%)</td>
<td>296 (16.8%)</td>
<td>24 (27.6%)</td>
</tr>
<tr>
<td><strong>MRI</strong></td>
<td>Chemical shift present†</td>
<td>0</td>
<td>1 (1.5%)</td>
<td>247 (14.0%)</td>
</tr>
<tr>
<td>Chemical shift absent‡</td>
<td>1 (1.0%)</td>
<td>8 (12.3%)</td>
<td>65 (3.7%)</td>
<td>15 (17.2%)</td>
</tr>
<tr>
<td><strong>Not done</strong></td>
<td>97 (99.0%)</td>
<td>56 (86.2%)</td>
<td>1455 (82.3%)</td>
<td>67 (77.0%)</td>
</tr>
<tr>
<td><strong>¹⁸F-FDG PET</strong></td>
<td>Low uptake†</td>
<td>0</td>
<td>0</td>
<td>123 (7.0%)</td>
</tr>
<tr>
<td>High uptake‡</td>
<td>2 (2.0%)</td>
<td>17 (26.2%)</td>
<td>4 (0.2%)</td>
<td>5 (5.7%)</td>
</tr>
<tr>
<td><strong>Not done</strong></td>
<td>96 (98.0%)</td>
<td>48 (73.8%)</td>
<td>1640 (92.8%)</td>
<td>72 (82.8%)</td>
</tr>
<tr>
<td><strong>CT absolute contrast washout</strong></td>
<td>s60%</td>
<td>0</td>
<td>0</td>
<td>19 (1.1%)</td>
</tr>
<tr>
<td>&lt;60%</td>
<td>0</td>
<td>0</td>
<td>2 (0.1%)</td>
<td>0</td>
</tr>
<tr>
<td><strong>Not done</strong></td>
<td>98 (100.0%)</td>
<td>65 (100.0%)</td>
<td>1746 (98.8%)</td>
<td>87 (100.0%)</td>
</tr>
<tr>
<td><strong>Maximum diameter on CT after ≥6 months</strong></td>
<td>Stable or reduced†</td>
<td>0</td>
<td>5 (7.7%)</td>
<td>611 (34.6%)</td>
</tr>
<tr>
<td>Increased‡</td>
<td>0</td>
<td>8 (12.3%)</td>
<td>6 (0.3%)</td>
<td>8 (9.2%)</td>
</tr>
<tr>
<td><strong>Not done</strong></td>
<td>98 (100.0%)</td>
<td>52 (80.0%)</td>
<td>1150 (65.1%)</td>
<td>53 (60.9%)</td>
</tr>
</tbody>
</table>

Data are n (%) or median (IQR). ACC = adrenocortical carcinoma. ACA = adrenocortical adenoma. HU = Hounsfield units. *Non-incidental modes of discovery included adrenal masses discovered after imaging done for either clinical signs and symptoms indicative of possible steroid excess or suggestive of a tumour (ie, mass effect with abdominal discomfort or symptoms with weight loss, low grade fever, or both). †Negative for ACC. Positive for ACC. HU could not be reliably measured and the tumour was classified as positive.
based on urine steroid metabolomics was consistently associated with lower false-positive rates (appendix pp 26–28). Imaging performance was generally better in non-ACC tumours smaller than 4 cm than in those 4 cm or larger. Unenhanced CT had a higher true-positive rate for ACC than a urine steroid metabolomics result indicating high risk of ACC (appendix p 26); MRI or ¹⁸F-FDG PET were used to assess very few ACCs (appendix pp 27–28), precluding comparisons with these methods.

247 (12·2%) of 2017 participants had positive results for tumour diameter and imaging characteristics. Among these, 95 of the 98 ACCs were true positives, leaving 152 false-positive results (figure 3). Participants with a urine steroid metabolomics result indicating a high risk of ACC and a positive result for either tumour diameter or imaging characteristics increased the PPV for ACC compared with having combined positive tumour diameter and imaging characteristic results (table 2, figure 3, appendix p 29). When urine steroid metabolomics was combined with tumour diameter, only four participants with ACC were dismissed, and when combined with imaging characteristics only three were dismissed (table 2).

In the triple testing strategy, a result of tumour diameter larger than 4 cm, positive imaging characteristics (attenuation >20 HU), and a urine steroid metabolomics result indicating high risk of ACC, yielded a group of 106 participants, including 81 of the 98 participants with ACC (figure 3), giving a PPV for ACC of 76·4% (table 2). The triple testing strategy classified 70 participants as having moderate risk of ACC, including 12 of the 98 participants with ACC (PPV 17·1%, table 2). The ability of the triple testing strategy to rule out ACC was high (table 2).

In a sensitivity analysis, when participants with mixed or aberrant steroid excess (n=42), clinical presentation of large adrenal mass with extra-adrenal metastases (n=13), and bilateral macronodular adrenal hyperplasia and isolated cortisol excess (n=22) were excluded, the high accuracy of urine steroid metabolomics compared with imaging was confirmed, as was the improvement of accuracy when urine steroid metabolomics was used in combination with imaging modalities (appendix pp 16, 30–32).

Among the 65 non-ACC malignant tumours, 46 (70·8%) had a tumour diameter greater than 4 cm, 63 (97·0%) had positive imaging characteristics, and seven (10·8%) had a urine steroid metabolomics profile indicating high risk of ACC. These tumours could not be reliably differentiated from ACA and other benign tumours with any of the combined testing strategies (appendix pp 17, 33–35).

**Discussion**

Our results have validated the diagnostic utility of urine steroid metabolomics in detecting ACC in participants with newly diagnosed adrenal masses. Diagnostic accuracy of urine steroid metabolomics was high compared with maximum tumour diameter and imaging characteristics, and the best performance was seen when these three methods were combined. We also showed that using a cutoff of 20 HU for unenhanced CT tumour attenuation increases the accuracy of imaging characteristic assessment for exclusion of ACC compared with the currently recommended cutoff of 10 HU, which has immediate implications for clinical practice.

1686 (83·6%) of the adrenal tumours in our study cohort had been discovered incidentally, yielding to our knowledge the largest prospective cohort of participants with adrenal incidentaloma. We achieved this study size via a comprehensive, multicentre, non-selective approach, completing recruitment in 5·5 years. This cohort size compares favourably with the largest retrospective adrenal incidentaloma cohort, in which 1096 patients were identified over a 15-year period (1980–95). The distribution of underlying pathologies in our cohort was similar to those in retrospective studies, showing the representativeness of our cohort for clinical practice.

96 of 98 ACCs in our study were larger than 4 cm. This cutoff was previously suggested to be sensitive (93%)
for differentiating ACC from ACA, but to have poor specificity.\cite{10,11,15,22} Given the non-selective nature of recruitment in EURINE-ACT and the small number of ACCs less than 4 cm in diameter, growth velocity of these tumours seems to be very rapid, making early-stage detection unlikely. 1549 (76.8%) of participants in the study underwent unenhanced CT. Only one ACC had attenuation below 20 HU, giving this cutoff considerably improved specificity compared with the 10 HU cutoff recommended in the European guidelines.\cite{11,15−17} Applying cutoffs of 4 cm for maximum tumour diameter and 20 HU for unenhanced CT tumour attenuation would, therefore, help to avoid unnecessary imaging procedures and adrenalectomies.

<table>
<thead>
<tr>
<th>ACC (n=98)</th>
<th>Other malignant tumours (n=65)</th>
<th>All non-ACC tumours (n=1919)</th>
<th>ACA (n=1767)</th>
<th>Other benign tumours (n=87)</th>
<th>Total (n=2017)</th>
<th>Percentage of ACC cases (95% CI)</th>
<th>Percentage of non-ACC cases (95% CI)</th>
<th>Likelihood ratio (95% CI)</th>
<th>Post-test probability of ACC (per 100 participants with results)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 cm</td>
<td>96</td>
<td>46</td>
<td>392</td>
<td>296</td>
<td>50</td>
<td>488</td>
<td>98.0% (92.8–99.9)*</td>
<td>20.4% (18.6–22.3)</td>
<td>4.8 (4.4–5.3)</td>
</tr>
<tr>
<td>&lt;4 cm</td>
<td>2</td>
<td>19</td>
<td>1527</td>
<td>1471</td>
<td>37</td>
<td>1529</td>
<td>2.0% (0.2–7.2)</td>
<td>79.6% (77.7–81.4)†</td>
<td>0.03 (0.01–0.10)</td>
</tr>
<tr>
<td>Imaging characteristics‡</td>
<td>Positive</td>
<td>97</td>
<td>63</td>
<td>396</td>
<td>289</td>
<td>44</td>
<td>99.0% (94.4–100.0)*</td>
<td>20.6% (18.8–22.5)</td>
<td>4.8 (4.4–5.3)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>1</td>
<td>2</td>
<td>1523</td>
<td>1478</td>
<td>43</td>
<td>1.0% (0.0–5.6)</td>
<td>79.4% (77.5–81.2)†</td>
<td>0.01 (0.00–0.00)</td>
</tr>
<tr>
<td>Urine steroid metabolomics</td>
<td>High risk of ACC</td>
<td>83</td>
<td>7</td>
<td>157</td>
<td>143</td>
<td>7</td>
<td>84.7% (76.0–91.2)</td>
<td>8.2% (7.0–9.5)</td>
<td>10.4 (8.7–12.3)</td>
</tr>
<tr>
<td></td>
<td>Moderate risk of ACC</td>
<td>13</td>
<td>28</td>
<td>655</td>
<td>578</td>
<td>49</td>
<td>13.3% (7.3–21.6)</td>
<td>34.1% (32.0–36.3)</td>
<td>0.39 (0.23–0.65)</td>
</tr>
<tr>
<td></td>
<td>Low risk of ACC</td>
<td>2</td>
<td>30</td>
<td>1107</td>
<td>1046</td>
<td>31</td>
<td>2.0% (0.2–7.2)</td>
<td>57.7% (55.4–59.9)</td>
<td>0.04 (0.01–0.14)</td>
</tr>
<tr>
<td>Combined-test strategies</td>
<td>Tumour diameter and imaging characteristics‡</td>
<td>&gt;4 cm and positive</td>
<td>95</td>
<td>45</td>
<td>152</td>
<td>83</td>
<td>24</td>
<td>247</td>
<td>96.9% (91.3–99.4)*</td>
</tr>
<tr>
<td></td>
<td>&lt;4 cm, negative, or both</td>
<td>3</td>
<td>20</td>
<td>1767</td>
<td>1684</td>
<td>63</td>
<td>3.1% (0.6–8.7)</td>
<td>92.1% (90.8–93.2)†</td>
<td>0.03 (0.01–0.10)</td>
</tr>
<tr>
<td></td>
<td>Tumour diameter and urine steroid metabolomics</td>
<td>&gt;4 cm and high risk of ACC</td>
<td>82</td>
<td>7</td>
<td>46</td>
<td>33</td>
<td>6</td>
<td>128</td>
<td>83.7% (74.8–90.4)</td>
</tr>
<tr>
<td></td>
<td>&gt;4 cm and moderate risk of ACC</td>
<td>12</td>
<td>20</td>
<td>130</td>
<td>85</td>
<td>25</td>
<td>12.2% (6.5–20.4)</td>
<td>6.8% (5.7–8.0)</td>
<td>1.8 (1.0–3.2)</td>
</tr>
<tr>
<td></td>
<td>&lt;4 cm, low risk of ACC, or both</td>
<td>4</td>
<td>38</td>
<td>1743</td>
<td>1649</td>
<td>56</td>
<td>4.1% (1.1–10.1)</td>
<td>90.8% (89.4–92.1)</td>
<td>0.04 (0.02–0.12)</td>
</tr>
<tr>
<td></td>
<td>Imaging characteristics and urine steroid metabolomics</td>
<td>Positive and high risk of ACC</td>
<td>82</td>
<td>6</td>
<td>43</td>
<td>35</td>
<td>2</td>
<td>125</td>
<td>83.7% (74.8–90.4)</td>
</tr>
<tr>
<td></td>
<td>Positive and moderate risk of ACC</td>
<td>13</td>
<td>28</td>
<td>155</td>
<td>97</td>
<td>30</td>
<td>13.3% (7.3–21.6)</td>
<td>8.0% (6.9–9.4)</td>
<td>1.69 (1.00–2.87)</td>
</tr>
<tr>
<td></td>
<td>Negative, low risk of ACC, or both</td>
<td>3</td>
<td>31</td>
<td>1721</td>
<td>1635</td>
<td>55</td>
<td>3.1% (0.6–8.7)</td>
<td>89.7% (88.2–91.0)</td>
<td>0.03 (0.01–0.10)</td>
</tr>
<tr>
<td></td>
<td>Tumour diameter, imaging characteristics, and urine steroid metabolomics</td>
<td>&gt;4 cm, positive, and high risk of ACC</td>
<td>81</td>
<td>6</td>
<td>25</td>
<td>17</td>
<td>2</td>
<td>106</td>
<td>82.7% (73.7–89.6)</td>
</tr>
<tr>
<td></td>
<td>&gt;4 cm, positive, and moderate risk of ACC</td>
<td>12</td>
<td>20</td>
<td>58</td>
<td>23</td>
<td>15</td>
<td>12.2% (6.5–20.4)</td>
<td>3.0% (2.3–3.9)</td>
<td>4.1 (2.3–7.3)</td>
</tr>
<tr>
<td></td>
<td>&lt;4 cm, negative, low risk of ACC, or a combination</td>
<td>5</td>
<td>39</td>
<td>1836</td>
<td>1727</td>
<td>70</td>
<td>5.1% (1.7–11.5)</td>
<td>95.7% (94.7–96.5)</td>
<td>0.05 (0.02–0.13)</td>
</tr>
</tbody>
</table>

Data in columns one to six are numbers of participants. ACC=adrenocortical carcinoma. ACA=adrenocortical adenoma. *Sensitivity. †Specificity. ‡Positive was classified as unenhanced CT attenuation >20 Houndsfield units in homogeneous tumours or heterogeneous tumours precluding measurement of attenuation.

Table 2: Performance of tests and combination test strategies
PPV was greater with urine steroid metabolomics than with tumour diameter and imaging characteristics. Combination of all three approaches gave the best overall detection of ACC, even when ACC tumours potentially identifiable by other parameters, such as mixed or aberrant steroid patterns or clinical presentations, were excluded. Based on our triple test strategy in the whole study cohort (n=2017), following the initial scan that provided tumour size measurements, only 488 participants (24·2%) would have required further imaging. Additionally, participants with a urine steroid metabolomics result indicating high risk of ACC could have undergone surgery for ACC earlier and fewer unnecessary surgeries for benign tumours would have been done. For participants with urine steroid metabolomics results indicating moderate risk of ACC in the triple strategy, the PPV was 17%. Rather than blanket surgical removal, these participants could be managed individually, such as with detailed re-review of scans by a multidisciplinary team and consideration of biopsy. Although histopathology of an adrenal biopsy cannot differentiate between ACC and ACA, it can be informative in patients with other benign and malignant adrenocortical tumours, including metastases of extra-adrenal cancers.

Strengths of our study include its prospective, consecutive, and non-selective recruitment, and the relatively short time needed for recruitment (avoiding biases due to change in diagnostic technologies and standards). Our previous proof-of-principle study of urine steroid metabolomics for ACC detection involved urinary steroid profiling by gas chromatography-mass spectrometry, which is a low-throughput method requiring highly specialised expertise. In this study, we used the high-throughput LC–MS/MS approach, which is much more widely available. Although a similar multisteroid profiling method has been described elsewhere, our method includes a machine-learning algorithm that was trained on steroid data measured by LC–MS/MS in a retrospective cohort, precluding error due to non-standardisation of steroid profiling results. Other strengths of our study include the exclusion of participants undergoing imaging for cancer monitoring and the non-communication of urine steroid metabolomics results to clinical centres so as not to influence the usual diagnostic process.

Weaknesses of this study include its observational nature; notably, recruitment centres could use their preferred imaging modalities and, therefore, the numbers assessed by MRI or ¹⁸F-FDG PET were too low for comprehensive assessment. Participating centres were specialist adrenal tumour centres, but in non-specialised secondary care settings the proportions of large and malignant tumours might be lower. Histopathology was done only for participants who had surgery or biopsy, meaning that our reference standard was heterogeneous. However, it was of higher quality than most earlier diagnostic studies and is representative of clinical practice. We did not do a centralised pathology review, but as the study sites were high-volume specialist centres with established adrenal pathology expertise, all pathologists applied relevant multifactorial scoring systems for assessment of malignant potential in adrenal cortical neoplasms.

In conclusion, our findings suggest that use of urine steroid metabolomics in diagnostic pathways could improve detection of ACC. We recommend a combined testing strategy of assessment with unenhanced CT, with a tumour attenuation cutoff of 20 HU, and urine steroid
metabolomics. We anticipate that these changes would substantially lessen the burden on and morbidity in patients with benign adrenal tumours and suspicious imaging findings, and lead to reductions in health-care costs due to decreased numbers of imaging procedures, time to surgery in ACC, and numbers of unnecessary surgeries.

Contributors

Declarations of interest
WA is an inventor and MB is a contributor on a patent on the use of steroid profiling as a biomarker tool for the differential diagnosis of adrenal tumours (PCT/GB2010/000274). All other authors declare no competing interests.

Data sharing
We will consider sharing de-identified, individual participant-level data that underlie the results reported in this Article on request of a receipt detailing the study hypothesis and statistical analysis plan. All requests should be sent to the corresponding author. The corresponding author and lead investigators of this study will discuss all requests and make decisions about whether data sharing is appropriate based on the scientific rigour of the proposal. All applicants will be asked to sign a data access agreement.

Acknowledgments
This work was supported by the European Commission under the Seventh Framework Programme (FP7/2007–13, grant agreement 259735, ENSAT, CANCER, awarded to GE, MF, FB, and WA); the UK Medical Research Council (Clinical Research Training Fellowship WT101671 awarded to VC); the Wellcome Trust (Mayo Scholarship awarded to IB), and the Wellcome Trust–University of Birmingham (BRC-1215-20009). The views expressed are those of the authors and not necessarily those of the NIHR, the UK Department of Health and Social Care, or the NIH. University of Birmingham (BRC-1215-20009). The views expressed are those of the authors and not necessarily those of the NIHR, the UK Department of Health and Social Care, or the NIH.

References