Adrenal incidentalomas — a continuing management dilemma

R Nawar and D Aron

Division of Clinical and Molecular Endocrinology, Department of Medicine, Case Western Reserve University School of Medicine, Cleveland, Ohio, USA
VA Health Services Research and Development Center for Quality Improvement Research, Louis Stokes Department of Veterans Affairs Medical Center, 10701 East Boulevard, Cleveland, Ohio 44106, USA

Abstract

Adrenal incidentalomas (AI), adrenal tumors detected through an imaging procedure done for reasons unrelated to adrenal dysfunction, is becoming a common clinical problem with the more frequent utilization of different imaging techniques. Most such tumors are benign and hormonally inactive. A variety of diagnostic strategies have been developed to distinguish the latter; however, they are still controversial. Even after a commissioned systematic review of the literature and a state of the science conference sponsored by the National Institutes of Health, the optimal strategy for hormonal screening of a patient with AI is unknown, but we anticipate further refinements and major advances in the field. Surgery is the ultimate solution for the diagnostic-therapeutic dilemma of AI. Careful planning is required, and the learning curve which influences clinical decision making is especially relevant to immediate outcomes. The benefit of making a diagnosis of a clinically significant AI must be considered in the context of the patient's overall condition and preferences.

Introduction

The term ‘adrenal incidentaloma’ (AI) refers to an adrenal mass unexpectedly detected through an imaging procedure done for reasons a priori unrelated to adrenal dysfunction or suspected dysfunction (Bertherat et al. 2002). First described more than two decades ago (Geelhoed & Spiegel 1981, Geelhoed & Druy 1982, Prinz et al. 1982), AI has become a common clinical problem, one that poses a challenging management dilemma (Aron 2002). The challenge is to recognize and treat the small percentage of AI that do pose a significant risk, either because of their hormonal activity or because of their malignant histology, while leaving the rest alone. The latter are benign and hormonally inactive and neither pose a risk to a patient’s health nor warrant the risks of further diagnosis and treatment. As imaging techniques improve, we can expect to encounter an increasing number of AI. This issue has gained increasing attention (Mansmann et al. 2004), and several approaches have been recommended including strategies for hormonal screening, radiologic testing, and histopathologic examination. However, even a commissioned systematic review of the literature and a state of the science conference sponsored by the National Institutes of Health have failed to end the controversy (Lau et al. 2002, Grumbach M et al. 2003). In the words of the eighteenth-century British philosopher David Hume, 'it is impossible to separate the chance of good from the risk of ill'. Medical decisions inevitably take place under conditions of uncertainty.

The magnitude of the problem

It is difficult to know the true prevalence of this entity, because of varied definitions and variability in methods and circumstances of detection; that is, the reasons for the imaging study. In addition, population-based studies, as opposed to institution-based studies that are dependent upon referral patterns, are needed. Incidentaloma is not so much a disease entity as a finding that may or may not represent a disease (Lau et al. 2002). While the prevalence of clinically inapparent adrenal masses found at autopsy is
1.4–2.9% (Kokko et al. 1967, Abecassis et al. 1985, Lau et al. 2002), it ranges from about 0.1% for general health screening with ultrasound, to 0.4–1.9% among patients evaluated for nonendocrinologic complaints, to approximately 4.4% among patients who have a previous cancer diagnosis (Herrera et al. 1991, Lau et al. 2002). Advancing age is associated with increasing frequency of adrenal masses. The prevalence of adrenal masses identified at autopsy increases from <1% among individuals less than 30 years of age to about 7% in those 70 years of age or older (Kloos et al. 1995, Lau et al. 2002). Sixty percent of AI occur between the sixth and the eighth decade at a mean age of 56±12.9 years (Mantero & Arnaldi 2000). There appear to be no gender differences in prevalence from autopsy studies or general health examinations; the reported higher frequency of AI in women may be an artifact related to their higher likelihood of undergoing imaging procedures. Patients with clinical features associated with functioning adrenal lesions (e.g., primary hyperaldosteronism, pheochromocytoma, Cushing’s syndrome and virilization) are more likely to have adrenal tumors. Although in many cases these features are subtle and do not prompt early diagnosis, whether one should consider these lesions to be truly ‘incidental’ is a matter of opinion. Although different authors have used different definitions, the majority agree that it is a mass discovered by imaging performed for reasons unrelated to suspected adrenal pathology. A difference in exclusion criteria is an important issue. Does the known presence of extra-adrenal cancer itself constitute an appropriate exclusion criterion? What about the fact that the complaints for which the diagnostic imaging was performed, turned out later to be caused by the adrenal mass?

The differential diagnosis of an incidentally discovered mass is extensive (Table 1), but most are nonsecreting cortical adenomas. In a recent systematic review that combined studies using the broadest definitions, adenomas accounted for 41%, metastasis 19%, adrenocortical carcinoma 10%, myelolipoma 9% and pheochromocytoma 8%, with other usually benign lesions such as adrenal cysts comprising the remainder (Aron 2002, Lau et al. 2002). It is evident that this distribution will vary as other inclusion and exclusion criteria are applied; that is, metastasis becomes much less common when patients with known extra-adrenal cancer are excluded. The development of a regional, national or international registry that utilized uniform operational definitions would go a long way toward clarifying these differences. The size of the lesion affects the distribution of etiologies: larger tumors are more likely to be malignant than smaller ones (Vierhapper & Heinze 2005). Among lesions larger than 6 cm, adrenal carcinomas comprised 25% and metastasis 18%, while adenomas accounted for only 18% (Lau et al. 2002); for tumors under 4 cm, adrenal carcinomas comprised 2% and adenomas 65%; and for tumors of 4–6 cm, adrenocortical carcinoma constituted 6% (Lau et al. 2002). Bilateral adrenal masses are found in about 10–15% of cases. When masses are bilateral, several diagnoses are more likely, including metastatic disease, congenital adrenal hyperplasia, lymphoma, infection (e.g., tuberculosis, fungal), hemorrhage, adrenocorticotrophic hormone (ACTH)-dependent Cushing’s syndrome, pheochromocytoma, amyloidosis and infiltrative disease of the adrenal glands.

A variety of rare genetic syndromes, such as Beckwith–Wiedemann, Li–Fraumeni, multiple endocrine neoplasia (MEN) 1, Carney complex, and

---

**Table 1 Differential Diagnosis of an Incidentally Discovered Adrenal Mass from Aron DC, Kievit J. (2003). “Adrenal Incidentalomas.” In Endocrine Surgery. Schwarz (ed.) Marcel Dekker, NY**

<table>
<thead>
<tr>
<th>Adrenal cortical tumors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenoma*</td>
</tr>
<tr>
<td>Carcinoma*</td>
</tr>
<tr>
<td>Nodular hyperplasia*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adrenal medullary tumors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pheochromocytoma*</td>
</tr>
<tr>
<td>Ganglioneuroma/neuroblastoma*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other adrenal tumors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myelolipoma</td>
</tr>
<tr>
<td>Metastases</td>
</tr>
<tr>
<td>Miscellaneous e.g., hamartoma, teratoma, lipoma, hemagioma</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Infections, granulomas, infiltrations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abscess</td>
</tr>
<tr>
<td>Amyloidosis</td>
</tr>
<tr>
<td>Fungal infection, e.g., histoplasmosis, coccidiomycosis, blastomycosis, tuberculosis</td>
</tr>
<tr>
<td>Sarcoïdosis</td>
</tr>
<tr>
<td>Cytomegalovirus</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cysts and pseudocysts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parasitic</td>
</tr>
<tr>
<td>Endothelial</td>
</tr>
<tr>
<td>Degenerative adenomas</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Congenital adrenal hyperplasia*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemorrhage</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pseudoadrenal masses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Splenic, pancreatic, renal lesions</td>
</tr>
<tr>
<td>Vascular lesions (esp. aneurysms and tortuous splenic veins)</td>
</tr>
</tbody>
</table>

| Technical artifacts                 |

*Potentially functions.
McCune–Albright syndrome, predispose to adrenocortical tumors. Similarly, some other disorders, such as multiple endocrine neoplasia 2 (2A, 2B), Von Hippel–Lindau disease and neurofibromatosis type 1, predispose to adrenomedullary tumors. Some of the molecular targets for these syndromes have been isolated. For example, overexpression of the insulin-like growth factor II (IGF II) gene has been associated with the tumors of the Beckwith–Wiedemann syndrome. Similarly, in the autosomal dominant Li–Fraumeni syndrome, germline mutations in the p53 tumor suppressor gene contribute to the high cancer risk, although adrenocortical carcinoma, is a very unusual presentation of this rare syndrome. Abnormalities in the RET-2 proto-oncogene (Pacak et al. 2001) have been associated with pheochromocytoma in MEN 2. We anticipate that, with time, we will discover how these findings are relevant to the more typical sporadic AI (Koch et al. 2002b).

Imaging of adrenal incidentaloma

Adrenal incidentalomas are typically detected during ultrasonography, computed tomography (CT) scanning or magnetic resonance imaging (MRI). Sometimes these techniques permit a diagnosis without further evaluation. However, it is important to recognize that even when they have reasonable sensitivity and in the aggregate can distinguish among types of lesions, imaging results may not be definitive on any individual case. AI, particularly those involving the right adrenal, may be discovered by ultrasound, but the technique has little differential diagnostic utility (Suzuki et al. 2001). It is operator-dependent and provides little information about the malignant potential of a mass, other than that based merely on diameter and solidity, and none about its functional status. However, ultrasound can be a simple, effective tool for follow-up. Rapid growth suggests malignancy (Fontana et al. 1999, Mansmann et al. 2004).

The CT scan is most commonly used in the assessment of AI. Adrenal adenomas are usually small, well-defined homogeneous lesions, with evident margins and large intracytoplasmic lipid content. The presence of hemorrhage, calcifications or necrosis is common but nonspecific. Large size, irregular shapes, vague contour, invasion into surrounding structures and high signal intensity usually denote malignancy (Angeli et al. 1997). However, size as such cannot be used to distinguish malignant from benign lesions with 100% accuracy (Lee et al. 1991, Singer et al. 1994, van Erkel et al. 1994, McNicholas et al. 1995, Sztolar & Kammerhuber 1998). Signal intensity can be very useful. High signal intensity can be expressed in either Hounsfield units (HU) (with an intensity higher than 10 or 20 units being used as diagnostic threshold) or in signal-lesion (SL) to signal-fat (SF) ratio (with SL/SF ratio above 1.5 suggesting malignancy). The threshold used determines the sensitivity and specificity of the test. At low threshold, sensitivity for malignancy is around 0.9 and specificity around 0.4 (Kievit & Haak 2000). Higher thresholds lead to better specificity (around 0.9), but with lower sensitivity (around 0.55) (Hamrahian et al. 2005). More recently, excellent results have been observed for the identification of adenomas by 10–15-min-delayed enhanced CT, since that adenomas are characterized by rapid washout of IV contrast. Using this method, a washout of 40–50% is highly suggestive of a benign mass with sensitivity and specificity of 96% and 100% respectively, whereas low washout percentages strongly suggest metastasis (for an overall accuracy of 96%) (Korobkin et al. 1998, Boland et al. 1999, 1998, Boland et al. 1997, 1998, Caoli et al. 2000, 2002).

The accuracy of MRI to the differentiation between benign and malignant tumors is comparable to that of CT scan. Usually, malignant masses are hypointense on T1-weighted images and hyperintense on T2-weighted images, with strong enhancement after contrast injection and delayed washout. More recent studies have used the lipid content of adenomas, which causes a loss in signal intensity on chemical-shift MRI (Outwater et al. 1995, 1996). However, MRI may be helpful in the diagnosis of pheochromocytoma, where it outperforms CT scanning. High signal intensity on T2-weighted MRI is suggestive of pheochromocytoma and originally was claimed to have nearly 100% accuracy (van Gils et al. 1991, van Erkel et al. 1994).

Experience has shown that adenomas are characterized by rapid washout of IV contrast. Using this method, a washout of 40–50% is highly suggestive of a benign mass with sensitivity and specificity of 96% and 100% respectively, whereas low washout percentages strongly suggest metastasis (for an overall accuracy of 96%) (Korobkin et al. 1998, Boland et al. 1999, 1998, Boland et al. 1997, 1998, Caoli et al. 2000, 2002). By combining data from several studies, sensitivity and specificity of T2-weighted MRI for pheochromocytoma were estimated at approximately 92% and 88% respectively (Kievit & Haak 2000). If such accuracy were indeed true, this would make T2-weighted MRI more accurate for diagnosing pheochromocytoma than some of the biochemical assessments.

Functional evaluation

When confronting an AI for which the diagnosis is not certain, one must address the adverse outcomes by which the patient can potentially be harmed: morbidity or mortality from hormonal excess or cancer and the anxiety that comes from knowing about a ‘tumor’ which might cause problems in the future (Aron 2001,
Hypersecretion of glucocorticoids, mineralocorticoids, sex steroids or catecholamines can produce clinical syndromes, each associated with morbidity and premature mortality. When a hormonal disorder is suspected, clinically appropriate, i.e., targeted, diagnostic testing can proceed apace. Controversy arises in the approach to screening for more subtle forms of hormonal excess (Aron & Kievit 2003, Mansmann et al. 2004). However, one should not infer the absence of endocrine activity by the AI solely because of the absence of clinically recognizable signs and symptoms. In fact, subclinical hormonal hypersecretion may be associated with or be a marker for premature morbidity and mortality. Therefore, part of the diagnostic evaluation should aim at assessing the presence of hormonal hypersecretion, specifically cortisol excess; catecholamine excess; and, if the patient is hypertensive, mineralocorticoid excess (Grumbach et al. 2003). In fact, virtually all diagnostic algorithms are variations on this theme (Copeland 1983, Ross & Aron 1990, Young 2000, Grumbach et al. 2003). An algorithm based on the 2002 NIH State-of-the-Science Conference is shown in Fig. 1 (Grumbach et al. 2003). This measured approach involves, in addition to careful clinical assessment, the biochemical screening of all patients for autonomous cortisol production and for pheochromocytoma and screening of all hypertensive patients for primary hyperaldosteronism. In addition, a size-based criterion for surgical removal of nonfunctioning lesions is recommended. The rationale for this approach follows.

**Adrenal pheochromocytomas**

Pheochromocytomas are a frequent cause of clinically silent adrenal masses; they constitute about 8% of patients with AI. Among patients referred to the Mayo Clinic for management of adrenal pheochromocytoma, 10% were discovered in the course of evaluation for AI (Young 2000). The classic features of pheochromocytoma are well known (headache, palpitations, diaphoresis, anxiety and sustained or paroxysmal hypertension), but most patients with AI proven later to harbor a pheochromocytoma lack those features, and the diagnosis is readily missed or delayed. In one

---

**Figure 1** Strategic approach from the NIH State-of-the-Science Conference (Grumbach et al. 2003).
study, there was a mean interval of 42 months between initial presentation and diagnosis; The delay was as long as 30 years (Mannelli et al. 1999). These ‘silent’ pheochromocytomas may also be lethal.

Clinical diagnosis of pheochromocytoma may be difficult. Even if careful history does not reveal the classic triad (headache, palpitations and diaphoresis) and even if physical examination demonstrates normotension, the diagnosis of pheochromocytoma cannot be excluded, with certainty. Positive and negative likelihood ratios are 40 and 0.36 respectively, increasing the probability of a pheochromocytoma from 3.5% to around 60% in case of a positive triad, and reducing it to 1% in case of its absence (and changing from a prior probability of 8%, as reported by Lau et al. (2002), to either 78% or 3%) (Kievit & Haak 2000). Hypertension provides even less reliable information, with positive and negative likelihood ratios of 2 and 0.6, changing the probability of a pheochromocytoma to 75% or 2% in the presence or absence of hypertension (or, again, from 8% to 15% and 5% respectively). Hormonal testing for urinary or plasma metanephrines count among the most reliable tests, with both sensitivity and specificity of around 95% (Lenders et al. 2002). Testing for urinary catecholamines and vanillymandelic acid (VMA) is less reliable, because of dangerously low test sensitivities of 80–85% (Mantero et al. 1997). Given the potential usefulness of T2-weighted MRI, a practical approach for ruling out pheochromocytoma would be a normal plasma or, if such testing were not available, urine metanephrine test, especially when combined with a negative T2-weighted MRI or positive testing for cortical hormonal hyperactivity.

Cortisol-secreting masses

Subclinical autonomous glucocorticoid hypersecretion (SAGH) has been found in 5–47% of patients with AI, in several studies using different study protocols and diagnostic criteria (Reincke et al. 1992, Angeli et al. 1997, Terzolo et al. 1998, Barzon & Boscaro 2000, Reincke 2000, Rossi et al. 2000). This autonomous hypersecretion is subtle and may be transient. These wide differences in prevalence are due to, among other things, lack of standardized criteria for diagnosis (Lau et al. 2002). These patients do not show clinical pattern of overt Cushing’s syndrome. Rather, they exhibit abnormal regulation of the hypothalamic-pituitary-adrenal (HPA) axis. They also have a high prevalence of obesity, hypertension, diabetes and insulin resistance. Abnormalities in bone turnover and bone mass have been reported as well (Torlontano et al. 1999, Tauchmanova et al. 2001, Francucci et al. 2002). In a recent study, involving 70 women with AI and 84 controls, evaluated by qualitative CT, the prevalence of vertebral fractures in pre- and postmenopausal AI patients, was significantly higher than in controls (Chiodini et al. 2004). Tauchmanova et al. (2002) used the criteria of the National Italian Group on Adrenal Tumors (AI-SIE) for SAGH: absence of clinical signs of cortisol excess and two abnormalities in the regulation of the HPA axis; failure to suppress serum cortisol to less than 83 nmol/l (3 μg/dl) by 2 mg dexamethasone; and the combination of a low-ACTH, high urinary free cortisol and blunted response to corticotropin-releasing hormone. In a study involving 28 consecutive patients and 100 age-, sex-, and body-mass index-matched controls, blood pressures (systolic and diastolic) were higher in patients than in controls. Similarly, higher levels of fasting glucose, insulin, total cholesterol and triglycerides were found in the patients with AI. In addition to increased insulin resistance, the AI patients had increased waist-to-hip ratios and high frequencies of hypertension (61%), lipid abnormalities (71%), and impaired glucose tolerance or diabetes (64%); 85% had multiple cardiovascular risk factors. Evidence of cardiovascular disease was present in a high proportion of patients, based not only on clinical evidence, but also on electrocardiogram, or carotid ultrasound examination. The degree to which these findings have impact on long-term morbidity of patients with SAGH remains to be determined. However, these metabolic abnormalities may improve after removal of the AI (Emral et al. 2003).

The transition from normal cortisol-ACTH feedback to completely autonomous cortisol production with cortisol hypersecretion is a continuum, but the point on this continuum that produces clinical morbidity is not clear. Tsagarakis et al. (1998) performed the dexamethasone suppression test in 61 patients with incidentally discovered adrenal masses. In a post-hoc analysis, patients were divided into three groups: patients with a post-dexamethasone level of cortisol of >70 nmol/l (group A, n=19), 30–70 nmol/l (group B, n=27) and <30 nmol/l (group C, n=15). Group A patients had significantly higher cholesterol and triglycerides concentrations than group C patients. In addition, the natural course of patients with SAGH is unknown. Of particular concern is the risk of progression to overt Cushing’s syndrome. Barzon and colleagues followed 130 nonoperated patients for at least 1 year. The cumulative risk for a nonsecreting adrenal incidentaloma to develop subclinical hyperfunction was 3.8% after 1 year and 6.6% after 5 years.
Of the 122 patients without subclinical autonomous cortisol secretion at initial AI diagnosis, three developed overt Cushing’s syndrome after 1–3 years (Barzon et al. 1999, 2002, Barzon & Boscaro 2000). In addition, one patient with autonomous cortisol secretion at initial diagnosis developed overt Cushing’s syndrome. Thus, for patients with masses with SAGH, estimated cumulative risk to develop overt Cushing’s syndrome was 11% after 1 year and 26% after 5 years.

**SAGH**

Diagnostic assessment of adrenocortical function is more easily resolved than that of adrenomedullary function, with the dexamethasone suppression test being the standard evaluation for subclinical autonomous cortisol hypersecretion; because pituitary ACTH-dependent hypercortisolism is not expected, the issue of false-negative tests is of much less concern (Terzolo et al. 1998). With dexamethasone suppression testing, a suppressed plasma cortisol at 8 am after 1 mg (or more) of dexamethasone rules out a cortisol-secreting adrenal tumor. The criterion for suppression is dependent upon the cortisol assay, but, typically, a value of <3 μg/dl would be considered adequate suppression. Inadequate suppression should be followed up with other tests: high-dose (8 mg) dexamethasone suppression test, high midnight cortisol levels and suppressed morning ACTH levels (should be measured at the same time as cortisol). The urinary free cortisol may be normal or slightly elevated and is thus less useful diagnostically. Salivary cortisol, although known to reflect plasma free cortisol better than the total plasma cortisol levels, has not yet been evaluated for SAGH (Findling & Raff 2001).

**Aldosteronoma**

The prevalence of mineralocorticoid-secreting mass in patients with AI is estimated at 1.6–3.8% (Barzon et al. 1998, Murai et al. 1999, Mantero et al. 2000). Hypertension with spontaneous hypokalemia (<3.5 nmol/l) was considered to be the hallmark of this entity, but normokalemic patients with primary hyperaldosteronism appear at a frequency 7–38% higher than previously thought (Stowasser 2001, Stowasser et al. 2003). Bernini et al. (2002), in their study, found primary hyperaldosteronism in 4% of patients who had an AI and were normokalemic, and in 5.5% of those with AI and hypertension. For those AI patients with hypertension, serum potassium and a plasma aldosterone concentration to plasma renin activity ratio (PAC-PRA) should be measured. A PAC-PRA greater than 30, and a PAC over 0.5 nmol/l is highly suggestive of autonomous aldosterone production (Grumbach et al. 2003). Additional testing should be done for confirmation of positive screening test results.

**Malignancy**

The diagnosis of an adrenal malignancy presents another set of issues. Although the differentiation between primary adrenal tumors and metastasis from extra-adrenal origin in oncology patients can be readily done without resorting to major surgery, the differentiation between benign and malignant primary adrenal tumor is more problematic.

**Adrenocortical carcinoma**

The most feared diagnostic possibility for AI is adrenal cancer, because of its poor prognosis, with a mean survival of approximately 18 months and 5-year survival of approximately 16% (Icard et al. 1992, Boscaro et al. 1995). These tumors can be functional or nonfunctional, with functional tumors accounting for 26–94%; frequently, the steroids synthesized have low biologic activity (Wooten & King 1993, Wajchenberg et al. 2000). Hypercortisolism, which can lead to Cushing’s syndrome, or a mixed Cushing-virilizing syndrome, is more common than virilization by androgen-secreting tumors (Del Gaudio & Del Gaudio 1993, Wajchenberg et al. 2000). Estrogen-secreting tumors causing feminization are rare, as are aldosterone-secreting carcinomas. Fortunately, clinically diagnosed cases of primary adrenal carcinoma are rare; the prevalence of adrenal carcinoma in general is approximately 12/1000 000 (Copeland 1983). However, the relative frequency of adrenal cancer varies considerably among AI series: 4.2% in the whole AI-SIE series, but 25% in another study (Terzolo et al. 1997, Mantero et al. 2000). The reasons for this apparent contradiction are unclear, given the high frequency of adrenal masses. However, in some series, the prevalence was determined by surgical findings. Some patients in such series may have been more likely to undergo surgery because of other clinical findings that are associated with adrenal cancer.

**Metastasis**

The adrenal glands are common sites of metastasis from extra-adrenal malignancy because of their high vascularity. Although adrenal metastases are typically bilateral and larger than 3 cm, they may be unilateral and small. In patients with a history of malignant
disease, metastases are the most common cause of an incidental adrenal mass, regardless of size, accounting for 50–75% of all incidentalomas in these patients (Belldegrun et al. 1986, Gillams et al. 1992, Liu et al. 2001). Carcinomas of the lung, breast, kidney and gastrointestinal tract and melanoma or lymphoma constitute the most common sources of adrenal metastases (Lau et al. 2002). In fact, adrenal metastases are found in 25–75% of those who die of epithelial malignancies. The source of the primary malignancy usually is known, or widespread disease is apparent, when AI is discovered. Metastatic cancer presenting as an isolated true AI is distinctly unusual. CT-guided adrenal biopsy is most useful in the diagnosis of adrenal metastases in patients with known extra-adrenal primary malignancies. Fine-needle aspiration is much less accurate in differentiating primary adrenal adenomas from adrenal carcinomas. Importantly, the possibility of pheochromocytoma should be assessed beforehand to avoid a potentially life-threatening hypertensive crisis.

**Issues in biochemical diagnosis**

Algorithms exist and continue to evolve for the diagnosis of hormonally active adrenal lesions. When patients present with signs or symptoms suggestive of a specific disorder (Cushing’s syndrome, Conn’s syndrome, pheochromocytoma or congenital adrenal hyperplasia), diagnostic evaluation can and should proceed apace. However, in the absence of clear findings, i.e., when the pretest probability of a particular disorder is low, biochemical evaluation is more challenging. Diagnostic test performance characteristics are usually determined for patients with clinically apparent disease. Inevitably, these tests will be less accurate when applied to patients lacking clinical fractures. Test sensitivity is likely to be lower than in study population from which the original characteristics were derived (spectrum bias); test specificity is also likely to be lower. Moreover, predictive value is dependent upon the prevalence of disease. Even a test with high sensitivity and specificity will, when used to detect a rare condition, falsely identify many nonaffected individuals as having the disease. These limitations notwithstanding, test performance in AI patients has been evaluated in a few studies (Lau et al. 2002). For example, two studies using unilateral uptake on NP-59 (radioiodinated cholesterol) scintigraphy as the reference standard indicator of autonomous activity of the incidentaloma (see below) found that no single test available discriminated well between unilateral and bilateral uptake on scintigraphy. Bardet et al. (1996) found that low 8 am dehydroepiandrosterone sulfate, low 8 am ACTH, basal urinary free cortisol, unsuppressed 8 am serum cortisol after overnight 1 mg dexamethasone suppression, unsuppressed day 2 serum cortisol, and unsuppressed day 2 urinary free cortisol after low-dose dexamethasone suppression had low sensitivity (0–50%) and moderate specificity (79–94%) to differentiate unilateral from bilateral scintigraphy uptake. In contrast, Valli et al. (2001) found that unsuppressed 8 am cortisol after overnight dexamethasone suppression had 100% sensitivity and 67% specificity to predict unilateral uptake. Studies of diagnosis of catecholamine excess and primary hyperaldosteronism reveal similar problems in sensitivity and specificity. The optimal strategy for hormonal evaluation of a patient with an incidentally discovered adrenal mass is unclear and remains controversial. Regardless of the approach used, it should be tailored to what is available; tests should be chosen based on their performance characteristics in the laboratories available to the clinicians ordering the test.

**Future directions in biochemical diagnosis involve the use of serum molecular markers**

A variety of tissue molecular markers, including mutant p53, the proliferation-associated antigen ki67 protein, and loss of heterozygosity at the 17p13 and 11p15 loci, have been used to distinguish malignant from benign adrenal tumors, or to differentiation between tumors of adrenocortical and adrenomedullary origin (Gicquel et al. 2001, Koch et al. 2002b). Overexpression of the insulin-like growth factor (IGF) II gene and IGF-binding protein 2 have been reported. Boulle et al. (2002) found no significant difference in plasma IGFBP-2 concentration between healthy controls and patients with complete remission or localized tumors. In contrast, patients with metastatic disease had significantly higher IGFBP-2 plasma levels than the control group (P < 0.001). IGFBP-2 levels in patients with metastatic disease were inversely correlated with survival (R² = 0.308; P = 0.0026). However, the overall sensitivity and specificity of this test were low; five patients (17.8%) with metastatic tumors had normal IGFBP-2 levels, and two patients (13.3%) in complete remission had high plasma IGFBP-2 levels, limiting its value in diagnosis and follow-up of adrenocortical carcinoma. Other markers, including survivin, an apoptosis inhibitor and a novel neuroendocrine marker for chromaffin cell tumor, have not
been reliable for distinguishing between benign and malignant tumors (Kolomecki et al. 2001a,b, Koch et al. 2002a). Even more recently, in a prospective clinical study, Reznik et al. (2004) were able to detect aberrant membrane receptors by in vivo stimulation tests in autonomously functioning unilateral adrenocortical adenomas, and they concluded that these receptors may be involved in the modulation of cortisol secretion in AI, with potential therapeutic consequences for the control of subclinical cortisol hypersecretion. However, the inconsistent results and relatively poor test characteristics argue against the routine clinical use of such markers at present, albeit with a few notable exceptions such as RET-2.

Another future direction involves function-based imaging methods

The two most frequently used function-based isotope imaging studies are $^{131}$I- or $^{123}$I-labeled metaiodobenzylguanidine (MIBG) and $^{133}$I-6-beta-iodomethyl-norcholesterol (NP-59) (Kurtaran et al. 2002). MIBG is mainly used to localize and identify pheochromocytoma, while NP-59 assesses adrenal cortical function as well as differentiates between benign and malignant adrenal tumors. MIBG has the advantage over CT or MRI in that it provides a whole-body image with the administration of one tracer dose. Pheochromocytomas can occur bilaterally and may not be confined to the periadrenal region (especially in malignant pheochromocytoma with metastasis). Sensitivity and specificity of MIBG are about 87% and 95% respectively (Kievit & Haak 2000). Positive and negative MIBG results, therefore, change the prior probability of pheochromocytoma in AI from 3.5% to around 40% and 0.5% respectively. The synthetic somatostatin analog $^{111}$In-octreotide seems less sensitive but is also able to visualize tumors that are undetected by MIBG scan (Tenenbaum et al. 1995).

NP-59 is taken up in functioning adrenocortical tissue and the presence of imaging activity concordant or discordant is determined (Kurtaran et al. 2002) by morphologic findings on CT or MRI. A concordant pattern is defined as a unilateral adrenal visualization at the site of the CT-detected mass. This is most consistent with hormonally active benign adenomas. A discordant pattern is with absent, decreased or distorted uptake by the adrenal mass, indicating adrenocortical carcinoma, metastasis or other non-functioning space-occupying adrenal lesion. However, experience with this technique has been limited to very few centers, and definitive conclusions about its usefulness in the assessment of AI cannot be drawn.

Newer techniques such as SPECT scanning are now being applied to the evaluation of AI (La Cava et al. 2003).

Assessment of strategies

AI typify the fact that incidental findings by their very nature pose a risk of overdiagnosis and overtreatment. Although most AI are of no significance beyond the anxiety they produce indirectly (not a trivial concern), some AI are clinically significant, and inadvertently leaving them alone might damage the patient’s health. Therefore, detection of an incidentaloma necessitates a conscious and conscientious decision regarding its management. Ideally, this decision/recommendation should be based on a careful weighing of the risks and benefits of each diagnostic and therapeutic step as well as individual patient preferences (Kievit & Haak 2000). The potential morbidity suggests a benefit of presymptomatic diagnosis. This benefit is conferred, however, more on an individual basis than at the population level; the health risk posed by AI, though real, is small because of the low prevalence of clinically significant tumors with hormonal activity or malignant potential, and the presence of a nonfunctional AI is relatively common.

All general approaches involve hormonal screening (Aron & Kievit 2003). However, they vary both in extent and in the specific screening tests. These differences appear to result from three issues: 1. differences in the use of empirical data to assess the probability of different disorders being present; 2. differences in assumptions about diagnostic and therapeutic effectiveness, again, traceable to available empirical evidence; and 3. differences in the weights that are, either implicitly or explicitly, being attributed to various outcomes. The first two issues relate to the quality and validity of the information used. Therefore, differences may be reduced if not eliminated by using standardized guidelines (such as the one advocated by the NIH panel), which are most evidence-based. The third issue is more difficult to address, because it relates to personal preferences, not evidence adduced from studies of others. Preferences may vary in the weight that is assigned to missing a relevant disorder, to short- or long-term morbidity and mortality, and to life expectancy, quality of life, etc. Even the term ‘relevant’ may be defined in different ways. For example, relevance may pertain to either ‘posing a severe risk to the patient’ or to ‘where outcome can be affected by diagnosis and treatment’. An adrenal metastasis is ‘relevant’ in the first sense, but may be considered less
relevant or even irrelevant in the second sense because of the impossibility to influence the bad prognosis. For such differences, there cannot be universal answers. Cost-effectiveness analysis may help to determine the best approach by explicitly outlining the alternatives.

Kievit and Haak performed a cost-effective analysis of 70 different strategies for the diagnosis and treatment of AI (Kievit & Haak 2000). In addition to the strategy used for comparison, ignoring the incidentaloma, they also analyzed strategies using one of eight single tests, various two-test sequences and sequences suggested by others. With respect to final outcomes, strategies differed strongly in costs (up to 10-fold) but only marginally in their health effects (up to 1.5%). The health risk of AI (in terms of potential loss of quality adjusted life years (QALYs)) mainly depended on characteristics of the patient and the incidentaloma. The choice of diagnostic-therapeutic strategy had far less impact, because two of the three significant disorders (adrenocortical cancer and metastasis) have a poor prognosis that is not drastically changed by treatment. There was no strategy that was clearly ideal. Strategies with low false-positive rates have higher false-negative rates. Studies with low false-negative rates cause more patients to receive unnecessary surgery. Moreover, later analysis showed that the factor to which the analysis was most sensitive was the degree of anxiety about having a mass without knowing what it is (personal communication). The evidence-based medicine movement has promoted transparency and accountability about the information. It is equally valuable to reveal how choices are guided by value judgments concerning process and outcome (Kievit & Haak 2000). The combination of best evidence and careful patient-centered value judgment is the key to good clinical practice for AI.

**Surgery for AI**

Surgery offers the ultimate solution to the diagnostic-therapeutic dilemma of AI. However, and sometimes despite an extensive evaluation, the diagnosis may be uncertain at the time of actual treatment, which makes surgery for AI different from surgery for a nonincidentaloma adrenal mass. Therefore, before surgery is undertaken, careful planning is required with respect to preoperative, intraoperative, and postoperative management (Aron & Kievit 2003). Preoperative management issues include decisions about hormonal blocking of a suspected pheochromocytoma or of adrenocortical hyperfunction. Concerning intraoperative care, either a laparoscopic or open surgical approach must be chosen, depending on both the size and the suspected nature of the incidentaloma. In addition, both surgeon and anesthetist may have to deal with the difficulties posed by a pheochromocytoma or by an intravascular invasion by the tumor (Aron & Kievit 2003). Postoperative care includes anticipating the consequences of the reduction in catecholamine levels that follow pheochromocytoma extirpation or the risk of secondary adrenal insufficiency after the removal of an AI with clinical or subclinical autonomous cortisol hypersecretion where of the HPA axis is suppressed (Reineke et al. 1992, Reincke 2000). Especially relevant to the immediate outcomes are the issues of the learning curve and the relationship between volume of procedures and outcome (Best et al. 2000, Slavin et al. 2000). For example, Fahlenkamp et al. (2002) described the experience with laparoscopic urologic surgery (including but not limited to adrenalectomy) at four German centers. There were 2407 such procedures. For the first 100 procedures, the complication rate was 13.3%, but for subsequent procedures, the rate was 3.6%. Similarly, Catarci et al. (2002) reported on 1006 cases and found a complication rate of 6% or the first 50 cases and 1.9% for the rest. The existence of a learning curve for these complex procedures should not be surprising, but it should also influence clinical decision making.

**Prognosis**

Of all the areas that could benefit from prospective studies, the issue of the natural history of AI stands out. Follow-up of patients with nonfunctioning adrenal masses suggests that the majority of adrenal lesions remain stable in size. About 5–25% increase in size by at least 1 cm and 3–4% decrease (Jockenhovel et al. 1992, Kasperlik-Zeluska et al. 1997, Barzon et al. 1999, 2002, Siren et al. 2000, Grossrubatscher et al. 2001, Lau et al. 2002, Libe et al. 2002). However, most studies are fairly small. The threshold for clinically significant increase in size is unknown, particularly since the reproducibility of size determination by imaging procedures is unknown. The risk of malignancy appears to be low. Up to 20% of patients develop hormone overproduction; while this less likely in tumors less than 3 cm, the risk a tumor’s becoming hormonally active appears to plateau after 3–4 years (Mansmann et al. 2004).

Little is known about whether early treatment is beneficial for these conditions before they cause significant symptoms, although the unpredictable nature of the pheochromocytoma — specifically, its ability to cause sudden death, and the insidious nature of the impact of the high cardiovascular risk.
R Nawar and D Aron: Adrenal incidentalomas — a continuing management dilemma

associated with SAGH — strongly suggest that early treatment would be beneficial. Moreover, the long-term prognosis of surgically treated benign adenomas causing catecholamine, cortisol or aldosterone excess appears to be reasonably good. Of note is the improvement in the cardiovascular risk factors, such as hypertension, hyperglycemia and hypercholesterolemia, after removal of an AI associated with subclinical autonomous cortisol hypersecretion. For patients found to have adrenal gland metastasis, prognosis is defined by the primary tumor’s histology, grade, stage and site. Approximately 25% of masses greater than 6 cm in diameter are adrenal cortical carcinomas, and these patients have very poor clinical outcomes. Among a large series of studies of adrenal cancer (usually not presenting as incidentalomas), 5-year survival was 19–62% with a median of 34% (Lau et al. 2002). There is some evidence suggesting that surgical extirpation of adrenal cancer at early stages may improve the survival rate. At more advanced stages, surgical debulking may increase the efficacy of adjuvant therapy (Kopf et al. 2001, Wajchenberg et al. 2000).

Summary and conclusions

The optimal strategy for hormonal screening of a patient with any incidentally discovered adrenal mass is unknown. Review of the endocrinologic literature supports the view that such patients are at somewhat increased risk for morbidity and mortality, indicating a benefit of early diagnosis for at least for some of the disorders. From a clinical perspective, our ability to determine accurately those at increased risk among the vast majority who are not at increased risk is poor. Given the limitations of diagnostic tests, effective hormonal screening requires a sufficiently high pretest probability to limit the number of false-positive results. This condition is met to varying degrees in the patient with an adrenal mass. Subjecting patients to unnecessary testing and treatment carries its own set of risks (Wolff & Kamerow 1990). Initial costs aside, testing may result in further expense and harm as false-positive results are pursued, producing the cascade effect described by Mold and Stein (1986) as a ‘chain of events (which) tends to proceed with increasing momentum, so that the further it progresses the more difficult it is to stop’. The extensive evaluations performed in some patients with incidentally discovered masses may reflect the unwillingness of many physicians and patients to accept uncertainty, even in the case of extremely unlikely diagnoses (Kassirer 1989). Not withstanding these considerations, and pending the accumulation of better evidence on which to base our decisions, the scheme outlined in the NIH consensus conference, modified according to patient’s overall condition and preferences, represents a prudent approach.

References

Best MA, Slavin LP & Aron DC 2000 To regionalize or not to regionalize: that is the question. II. Volume-outcome


R Nawar and D Aron: Adrenal incidentalomas — a continuing management dilemma


Korobkin M, Brodeur FJ, Francis IR, Quint LE, Dunick NR & Goodis M 1996 Delayed enhanced CT for differentiation of benign from malignant adrenal masses. Radiology 200 737–742.


OutwaterEK, Siegelman ES, Radecki PD, Piccoli CW & Mitchell DG 1995 Distinction between benign and


