Original research

Patient-specific workup of adrenal incidentalomas

Romy R. de Haan[a], Johannes B.R. Visser[b], Ewoud Pons[a,c], Richard A. Feelders[d], Uzay Kaymak[b], M.G. Myriam Hunink[e,f,g], Jacob J. Visser[a,*]

[a] Department of Radiology, Erasmus Medical Center, Rotterdam, The Netherlands
[b] Industrial Engineering & Innovation Sciences (IE & IS), Eindhoven University of Technology, Eindhoven, The Netherlands
[c] Department of Medical Informatics, Erasmus Medical Center, Rotterdam, The Netherlands
[d] Department of Internal Medicine, Division of Endocrinology, Erasmus Medical Center, Rotterdam, The Netherlands
[e] Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands
[f] Department of Radiology, Erasmus Medical Center, Rotterdam, The Netherlands
[g] Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, USA

ARTICLE INFO

ABSTRACT

Purpose: : To develop a clinical prediction model to predict a clinically relevant adrenal disorder for patients with adrenal incidentaloma.

Materials and methods: : This retrospective study is approved by the institutional review board, with waiver of informed consent. Natural language processing is used for filtering of adrenal incidentaloma cases in all thoracic and abdominal CT reports from 2010 till 2012. A total of 635 patients are identified. Stepwise logistic regression is used to construct the prediction model. The model predicts if a patient is at risk for malignancy or hormonal hyperfunction of the adrenal gland at the moment of initial presentation, thus generates a predicted probability for every individual patient. The prediction model is evaluated on its usefulness in clinical practice using decision curve analysis (DCA) based on different threshold probabilities. For patients whose predicted probability is lower than the predetermined threshold probability, further workup could be omitted.

Results: : A prediction model is successfully developed, with an area under the curve (AUC) of 0.78. Results of the DCA indicate that up to 11% of patients with an adrenal incidentaloma can be avoided from unnecessary workup, with a sensitivity of 100% and specificity of 11%.

Conclusion: : A prediction model can accurately predict if an adrenal incidentaloma patient is at risk for malignancy or hormonal hyperfunction of the adrenal gland based on initial imaging features and patient demographics. However, with most adrenal incidentalomas labeled as nonfunctional adrenocortical adenomas requiring no further treatment, it is likely that more patients could be omitting from unnecessary diagnostics.

1. Introduction

An adrenal incidentaloma is an adrenal mass detected on imaging studies performed for indications other than to evaluate the adrenal gland [1]. Advances in imaging and the widespread availability of imaging technology has resulted in detection of an increasing number of incidental findings [2]. The prevalence of adrenal incidentalomas found on computed tomography (CT) scans varies from 2.5% to 4% for abdominal CT, and 4.2% for thoracic CT, in adult populations [3–5]. This prevalence increases with age, up to 10% at an age of 70 or above [1,6–8]. As the population ages and the use of imaging technology intensifies, these incidentally discovered tumors would become a more prevalent diagnostic challenge [1,2,9].

Although most adrenal incidentalomas are nonfunctional

adrenocortical adenomas requiring no further treatment, diagnostic workup is needed to determine whether these incidental findings are malignant or cause hormonal hyperfunction [2,6]. According to current endocrine and surgical guidelines, every patient is adhered to the same expensive cascade of tests and procedures [10,11]. This results in significant cost. Berland et al. [20,10] recommend biochemical evaluation only in patients with clinical signs or symptoms of adrenal hyperfunction [11]. Furthermore, by evaluation of these incidental findings, patients are exposed to radiation from the CT, which increases the probability of cancer [9]. Each patient has a 1 in 430–2170 chance of developing cancer [9]. Cawood and colleagues showed that this potential risk of cancer is similar to that of the adrenal becoming malignant during the average recommended CT scan follow-up of 3 years.

It is important to identify all adrenal incidentalomas that are

malignant and/or cause hormonal hyperfunction, but at the same time avoiding patients from unnecessary workup. Therefore, clinical management of adrenal incidentaloma should be tailored to the individual patient. Personalized medicine will help produce more efficient and effective diagnoses and treatment, and will lead to better prognoses for patients at both the individual and population level. The availability of patient- and disease-related data in today's healthcare workforce is a significant resource in the development and application of this individual-based approach [12]. Therefore, by using these data preventive or therapeutic interventions can be concentrated on the patients who will benefit, and at the same time sparing expense and side effects for those who will not [13].

The purpose of this retrospective study was to develop a clinical prediction model to predict clinically relevant adrenal disorder for patients with adrenal incidentaloma.

2. Materials and methods

2.1. Patient identification

This retrospective study is approved by the institutional review board, with waiver of informed consent. For filtering of potential adrenal incidentaloma cases, natural language processing (NLP) is used in all thoracic and abdominal CT reports from 2010 till 2012 from a query of a searchable database of a major university-affiliated hospital. All patients with initial presentation or imaging workup for adrenal incidentaloma during this period are identified. Reports containing key sentences that mention the adrenal are selected. Non-relevant anatomic identifiers are removed from key sentences, and frequent negative patterns are subsequently filtered. A machine learning selection algorithm is trained using an annotated set of 500 key sentences and applied as final filter. By applying the NLP pipeline as proposed by Pons and colleagues [14], 139,329 radiology reports were removed from the dataset. Through this pipeline, we managed to increase finding prevalence of potential adrenal incidentaloma [14]. This corresponds with 2122 unique potential adrenal incidentaloma patients (Fig. 1).

2.2. Study population

All adult patients (≥18 years) with an adrenal mass incidentally found on CT are included in this study. Patients with recorded complaints possibly indicating adrenal disease at the moment of finding, patients with a vague described adrenal disorder presumably smaller than 1 cm in diameter, patients with expected metastasis, and patients with a history of metastatic disease are not considered to be incidentalomas and are therefore excluded from this study. Song et al., found no malignant mass among adrenal incidentalomas in low-risk patients, therefore patients with a history of malignancy are included [15]. Patients who were referred from another hospital are excluded, as no data regarding potential work- and follow-up is available in the electronic health record at our hospital. After re-measurements of the remaining incidentalomas, patients with an adrenal mass smaller than 1 cm are excluded.

2.3. Data collection

Patient workup data of all adrenal incidentaloma cases are collected from the electronic medical record by manual investigation of radiology reports, clinical letters and pathology reports. In addition, medication, lab tests and patient data are acquired from the respective hospital databases.

2.4. Data analysis: descriptive statistics

For patients included in the study demographical data, the radiological characteristics of the lesion, values of the biochemical

evaluation, and adrenal disorder are recorded. Because tumor size appeared to be lognormally distributed it is log-transformed. Surface of the nodule is calculated as the surface of an ellipse by using the dimensions of the long en short axis and laterality of the nodule. All patients outcome are verified with the nationwide network and registry of histo- and cytopathology in the Netherlands (PALGA).

2.5. Data analysis: predictive modeling

A clinical prediction model is developed to predict if an individual patient is at risk for a clinically relevant adrenal disorder. Stepwise logistic regression is used to construct the prediction model by using age and sex of the patient, the use of anti-hypertensive drugs, history of malignancy, surface, size and laterality of the nodule. For a total of 10 iterations a validation set is constructed based on a 15% hold-out sample with the original class distribution. The remaining data is used for training and tuning of the model using 10-fold cross-validation. The dataset is heavily imbalanced given that just 5% is clinically relevant. To overcome the problem of unbalanced data, the synthetic minority over-sampling technique (SMOTE) is used [16]. Hence a 50-50 class distribution of the target variable is generated by interpolation between existing cases. The prediction model predicts, at the presentation of an adrenal incidentaloma, if a patient is at risk for malignancy or hormonal hyperfunction of the adrenal gland. The model generates a predicted probability for every individual patient.

2.6. Model validation

For validation of the prediction model 10-fold cross-validation is used to better estimate the performance of the model. The performance metrics used are area under the curve (AUC), cohen's kappa and area under kappa (AUK). Cohen's kappa is chosen for two reasons: [1] it takes into account the agreement occurring by chance, and [2] it favors correct classification of the minority class over that of the majority class. These are properties not accounted for by AUC [17]. This is useful when dealing with an imbalanced dataset where it is more important to correctly predict the minority class. The performance of the model is evaluated based on the over-sampled test sets, using a threshold probability of 50%, as well as the validation dataset with original class distribution of the target variable.

2.7. Decision curve analysis

The prediction model is evaluated on its usefulness in clinical practice using decision curve analysis (DCA). DCA is a method for evaluating the benefits of a diagnostic test across a range of patients preferences for accepting risk of under- and overtreatment to facilitate decisions about test selections and use [18]. The prediction model generates a probability of having a clinically relevant adrenal disorder for every individual patient. Additional workup is needed to determine the diagnosis if the likelihood is near one. At some probability between 0 and 1, patients are unsure whether or not the be treated [18]. Threshold probability are determined by a panel of experts and are set on 1%, 1,5% and 2%. For patients whose predicted probability of the prediction model is lower than the predetermined threshold probability, diagnostic workup could be omitted.

3. Results

A total of 2122 potential adrenal incidentalomas are identified by using NLP. Of these, 1487 patients are excluded after manual inspection based on the inclusion and exclusion criteria (Fig. 1). Table 1 displays patient demographics and diagnostic workup features of the final cohort of 635 adrenal incidentaloma patients. Patients are predominantly older, averaging 63 years of age, with more females than males. Almost one-fifth of these patients (N = 118) used anti-hypertensive drugs at