Predicting serious complications in patients with cancer and pulmonary embolism using decision tree modelling: the EPIPHANY Index

A Carmona-Bayonas *,1, P Jiménez-Fonseca2, C Font3, F Fenoy4, R Otero5, C Beato6, J M Plasencia7, M Biosca8, M Sánchez9, M Benegas9, D Calvo-Temprano10, D Varona11, L Faez2, I de la Haba12, M Antonio13, O Madridano14, M P Solís2, A Ramchandani15, E Castañón16, P J Marchena17, M Martin14, F Ayala de la Peña1 and V Vicente1 on behalf of the Asociación de Investigación de la Enfermedad Tromboembólica de la Región de Murcia (the Region of Murcia’s Association of Thromboembolic Disease Research)

Background: Our objective was to develop a prognostic stratification tool that enables patients with cancer and pulmonary embolism (PE), whether incidental or symptomatic, to be classified according to the risk of serious complications within 15 days.

Methods: The sample comprised cases from a national registry of pulmonary thromboembolism in patients with cancer (1075 patients from 14 Spanish centres). Diagnosis was incidental in 53.5% of the events in this registry. The Exhaustive CHAID analysis was applied with 10-fold cross-validation to predict development of serious complications following PE diagnosis.

Results: About 208 patients (19.3%, 95% confidence interval (CI), 17.1–21.8%) developed a serious complication after PE diagnosis. The 15-day mortality rate was 10.1%, (95% CI, 8.4–12.1%). The decision tree detected six explanatory covariates: Hestia-like clinical decision rule (any risk criterion present vs none), Eastern Cooperative Group performance scale (ECOG-PS; \(<2 vs \geq 2\)), O₂ saturation (\(<90 vs \geq 90\%\)), presence of PE-specific symptoms, tumour response (progression, unknown, or not evaluated vs others), and primary tumour resection. Three risk classes were created (low, intermediate, and high risk). The risk of serious complications within 15 days increases according to the group: 1.6, 9.4, 30.6%; P<0.0001. Fifteen-day mortality rates also rise progressively in low-, intermediate-, and high-risk patients: 0.3, 6.1, and 17.1%; P<0.0001. The cross-validated risk estimate is 0.191 (s.e. = 0.012). The optimism-corrected area under the receiver operating characteristic curve is 0.779 (95% CI, 0.717–0.840).

Conclusions: We have developed and internally validated a prognostic index to predict serious complications with the potential to impact decision-making in patients with cancer and PE.

*Correspondence: Dr A Carmona-Bayonas; E-mail: alberto.carmonabayonas@gmail.com

Received 7 October 2016; revised 6 February 2017; accepted 7 February 2017; published online 7 March 2017

© 2017 Cancer Research UK. All rights reserved 0007 – 0920/17
Pulmonary embolism (PE) is one of the most common and feared complications in cancer patients, given its frequency and the suffering it entails (Sørensen et al., 2000). There is a 15–30% prevalence rate of PE in necroscopic series for the most thrombotogenic malignancies, owing to interactions between the mechanisms of tumorigenesis, haemostatic activation, and other factors (Swensen and Karwinski, 1989). The introduction of multidetector computed tomography (CT) has increased detection rates of incidental PE, present in 2–8% of the studies performed in cancer patients (Dentati et al., 2010). In some recent series, incidentally diagnosed PE accounted for a 50% of embolic events (Font et al., 2014, 2016). On the other end of the spectrum of severity, PE is also a common cause of fatal events in daily practice, as well as in trials with new targeted therapies (Ranpura et al., 2011; Den Exter et al., 2013).

Classifying prognosis in PE is important, in as much as episodes classified as low risk might be eligible for support reductions (e.g., outpatient management or early discharge, etc.), thereby lowering costs and enhancing patient comfort without compromising safety. In contrast, subjects at higher risk should receive stepped up care (inpatient management or early discharge, etc.), thereby lowering costs and enhancing patient comfort without compromising safety. Nevertheless, using any of them to assist in decision-making involves problems, not least of which is that their suitability for outpatient treatment on pragmatic clinical decision rules (CDR), such as the HESTIA study eligibility criteria, which are based prominently on altered vital signs and risk of bleeding (Siragusa et al., 2005; Zondag et al., 2011; Font et al., 2014; Weeda et al., 2016).

On the other hand, prognostic multivariate models have been created, such as the RIETE registry scale and POMPE-C score, that predict 30-day mortality probability following PE (Kline et al., 2012; Den Exter et al., 2013); although, at best, they are marginally superior to other classifications developed for PE in the general population (e.g., PESI or sPESI; Carmona-Bayonas et al., 2016). Nevertheless, using any of them to assist in decision-making involves problems, not least of which is that their suitability for incidental PE has yet to be proven (Kline et al., 2012; Den Exter et al., 2013). Furthermore, they are not sensitive to competitive risks, such as increased bleeding, responsible for some 10% of early mortality, or cancer progression, which accounts for 50% of 30-day mortality after a PE event (Den Exter et al., 2013; Carmona-Bayonas et al., 2016).

Consequently, there is no adequate prognostic stratification method for incidental and symptomatic PE. In this study, we have attempted to refine the classification of the entire spectrum of cancer-associated PE by combining an adaptation of the HESTIA criteria with other explanatory covariates and modelling a decision tree procedure.

**MATERIALS AND METHODS**

**Patients.** The source of information is an observational registry of consecutive cases of cancer-associated PE, who received care at 14 Spanish hospitals between 2004 and 2015 (Registro de Embolia Pulmonar en Pacientes con Neoplasias, EPHIPANY registry for its Spanish acronym). This registry’s design, methods, and characteristics have been previously reported in depth (Carmona-Bayonas et al., 2016; Font et al., 2016; Plasencia-Martinez et al., 2016). Briefly put, the basic eligibility criteria required that patients be adults (≥ 18 years) with a PE diagnosis confirmed by means of objective imaging (CT angiography scans, high probability scintigraphy, or CT scheduled to assess tumour response or for other reasons). In order to choose a truly oncological population, subjects were withdrawn from the study if the PE had occurred more than 1 month prior to the diagnosis of cancer, or if more than 1 month had elapsed since completing adjuvant chemotherapy. Patients were also excluded if they had not received anticoagulant therapy within 24 h of diagnosis or if they were already in the registry because of a previous PE episode.

**Study design.** The main objective of this study was to develop a prognostic model, the EPHIPANY index, for cancer patients and both incidental, as well as symptomatic PE. Given that it was a non-intervention database, the data reflect genuine clinical profiles and the decisions physicians make in line with their clinical practice. The data were collected from medical records or directly from the patients, together with clinicians with experience in cancer support treatment and radiologists who are subspecialised in diseases of the chest. All the investigators were trained in the study protocol requirements and the data were monitored in situ or by phone. The data were gathered by means of an electronic capture system, designed to refine inconsistencies and resolve data errors in real time. Data acquisition was not blinded. The minimum observation period was 3 months from the time PE was diagnosed, although longer follow-up was required whenever possible. The variables were collected during routine or unscheduled medical appointments.

**Variables.** The main outcome measure in this study was the occurrence of a serious medical condition between PE diagnosis on imaging and 15 days later. Serious complications are events that lead to serious clinical deterioration or death; for example, systolic blood pressure < 90 mm Hg, acute respiratory failure, right-side heart failure, acute kidney failure, major bleeding, or any other event the investigator deems serious (Supplementary Table 1).

Other secondary end points were all-cause 30-day mortality, the cause of 30-day mortality, and 30-day venous or arterial rethrombosis. 'Rethrombosis’ was defined as a second thrombotic event after appropriate PE treatment or progression of a previous venous thromboembolism (VTE) despite proper anticoagulant therapy. Rethrombosis was not considered to be a serious complication in the absence of the afore-named criteria. An autopsic diagnosis notwithstanding, researchers attributed the cause of death on the basis of a clinical history review and findings from complementary testing. Demise was deemed to be due exclusively to PE when there was a direct causal relationship through a concatenation of events associated to the thrombosis pathophysiology. ‘Mixed’ deaths were defined by the presence of a temporal relationship between patient demise and PE, although other intercurrent complications (e.g., infection or tumour progression) might plausibly play a relevant role. Death was considered unrelated to PE if there was no temporal relationship or concatenation of clear events. ‘Multiple’ was accepted as a response when there was a resumption of overlapping causes.

The potential explanatory covariates were selected after a bibliographic review and consultation with experts, taking into account their availability at patients’ bedside. Data recording did not allow for lost data for outcomes, survival times, and basic demographic and clinical characteristics (vital signs, tumour status, performance status, etc.).

The ‘CDR variable’ was defined as adaptation of Hestia’s study eligibility criteria used in previous studies (Zondag et al., 2011; Weeda et al., 2016). These criteria are typified by the presence of at
least one of the following: systolic blood pressure <100 mm Hg, arterial oxygen saturation <90%, respiratory rate >30 breaths per minute, pulse >110 beats per minute, sudden or progressive dyspnoea, other serious complications, constituting admission criteria in and of themselves, and clinically relevant bleeding, high risk of bleeding, or platelets <50 000 mm$^{-3}$. The CDR was assessed immediately prior to the time of radiological diagnosis of PE.

Other explanatory covariates included: age, gender, tumour stage, type of cancer, use of targeted cancer therapies, tumour response at the time of PE based on radiological criteria, Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 (Eisenhauer et al, 2009), Eastern Cooperative Group Performance Status scale (ECOG-PS; Oken et al, 1982), chronic obstructive pulmonary disease, prior cardiovascular disease, chronic kidney failure, concurrent deep vein thrombosis or a history of VTE, development of PE during treatment for a previous VTE, troponin levels (normal or high), creatinine clearance (normal, low), incidental or symptomatic diagnosis of the PE, presence of PE-specific symptoms, right ventricular diameter, additional radiological findings, Qandali index (Qandali et al, 2001), interventricular septal anomalies, presence of a single or multiple PE, oxygen saturation, blood pressure, heart and respiratory rates, previous tumour bleeding, prior use of antiaggregants, and major surgery in the previous month. Standardised definitions were used for each variable (Supplementary Table 1).

Development of a decision tree model. Classifications based on decision tree modelling seek to discover how the outcome variable is linked to the potential explanatory factors and, specifically, the configuration of these factors. This method is considered appropriate as the contribution of the explanatory covariates cannot be assumed to be necessarily additive or linear (Yohannes and Hoddinott, 1999; Lewis, 2000). The Exhaustive CHAID algorithm builds a decision tree by means of repeated partitions of each subset into two or more child nodes, beginning with the full database (Biggs et al, 1991). This methodology was used to determine the strength of association between the presence of serious complications within 15 days and the previously mentioned potential predictors. To determine the best split in each node, the categories of each predictor were merged into pairs until statistically significant differences were no longer observed within each component of the pair in comparison to the target variable. The predictors that produced the most significant partitions were then recursively chosen. Thus, the algorithm identified the main interactions and built subgroups defined by the different sets of independent variables. The level of significance for splitting nodes ($\chi^2$) was 0.05. The Bonferroni method was used to adjust the value of significance.

To cope with the overfitting and instability inherent to the decision tree, a 10-fold cross-validation procedure was applied. Thus, the data were randomly divided into 10 equal subsets. Trees were systematically built in 9 of those partitions (training subsets) and then tested in the remaining group (testing subset). Cross-validation produces a single, final model. 'Risk' is defined as the proportion of cases incorrectly classified by each of the individual trees; whereas the cross-validated risk estimate is the average of the proportion of cases incorrectly classified by each of the individual trees. The predictors that produced the most significant partitions were then recursively chosen. Thus, the algorithm identified the main interactions and built subgroups defined by the different sets of independent variables. The level of significance for splitting nodes ($\chi^2$) was 0.05. The Bonferroni method was used to adjust the value of significance. The percentage of serious complications was calculated in each of the terminal nodes of the tree and this was used to create the EPHIPANY rule, a simplified classification based on three categories: low, intermediate, and high risk. The predictive value of this classification was estimated by means of the odd’s ratio (OR) and its 95% confidence interval (95% CI) between the groups thereby generated and the appearance of the study end point. Cumulative hazards curves were calculated to establish the changes in hazard over time for each prognostic category.

RESULTS

Patient characteristics. Patient characteristics are summarised in Table 1. Pulmonary embolism was incidentally diagnosed in 53% of the cases. Twenty-eight percent ($n = 302$) of the episodes were
treated at home. All patients received anticoagulation (initial therapy with low-molecular weight heparin in 92%). At the time of PE diagnosis, 73.6% of the patients had a metastatic tumour and 53.6% were receiving chemotherapy. The most common tumours were breast, lung, and colon cancer, accounting for 54.7% of the series. The recruitment process is illustrated in Figure 1.

Outcomes. The main end point of this study, serious complications within 15 days, occurred in 208 patients (19.3%; 95% CI, 17.1–21.8%). The 15-day mortality rate was 10.1% (95% CI, 8.4–12.1) and of the 109 patients who died within that period, 45 (41%) did so as a result of tumour progression and not PE (Table 2). The rates of embolic recurrence and major bleeding were 4% and 2%, respectively.

Decision tree. Figure 2 shows the decision tree model with the 15-day, serious complications data for each end node. The Exhaustive CHAID method selected six explanatory covariates from the initial 39: the Hestia-like CDR variable (any risk factor present vs none), ECOG-PS (<2 vs ≥2), oxygen saturation (<90 vs ≥90%), presence of PE-specific symptoms, previous tumour response evaluation (tumour progression, unknown, or not evaluated vs others), and prior surgical resection of primary tumour. While other additional nodes involving more variables could be generated, they did not provide any incremental risk discrimination.

The best predictor in the root node was the Hestia-like CDR variable; the episodes that did not meet any of these criteria had a lower risk of serious complications within 15 days, in comparison with episodes that satisfied at least one of them (4.7 vs 29.7%; OR 0.11, 95% CI, 0.07–0.18; P < 0.0001) and 15-day mortality (2.5 vs 15.6%; OR 0.13, 95% CI, 0.07–0.25; P < 0.0001), respectively. The decision tree makes it possible to elaborate on the prognostic stratification in seven terminal nodes. For purposes of practicality, they are summarised into three risk categories: high, intermediate, and low. Supplementary Table 2 outlines the demographic and clinical characteristics of each subgroup.

Low risk encompasses patients without any Hestia-like CDR criteria, and with controlled tumours or resected primary tumours, with a risk of serious complications of between 1.4–3.4% and 0.3% 15-day mortality. Tumours with any of the CDR risk factors were at high risk (complications rate, 20–55%), with the exception of the group consisting of patients with a good performance status, and no PE-specific symptoms, who had an intermediate level of risk. This risk group would also include all stable patients having uncontrolled or unevaluated tumours, and without surgery for the primary tumour, with a 10.6% risk of complications. The cross-validated risk estimate is 0.191 (s.e. = 0.012); the optimism-

---

**Table 2. Main outcomes**

| Outcome                                      | Overall data set | %, (95% CI) |
|----------------------------------------------|------------------|-------------|
| 15-day serious complications<sup>a</sup>     | 208              | 19.3%, (17.1–21.8) |
| Systolic BP < 90 mm Hg                       | 50               | 4.7%, (3.5–6.1)          |
| Acute respiratory failure                    | 86               | 8.0%, (6.4–9.8)          |
| Fibrinolysis                                 | 7                | 0.7%, (0.2–1.4)          |
| Major bleeding                               | 54               | 5.0%, (3.8–6.5)          |
| Acute right ventricular failure              | 22               | 2.0%, (1.3–3.1)          |
| Acute renal failure                          | 17               | 1.6%, (0.9–2.5)          |
| Admission to ICU                             | 21               | 2.0%, (1.2–3.0)          |
| Cardiopulmonary resuscitation                | 4                | 0.4%, (0.1–0.9)          |
| Non-invasive ventilation                     | 9                | 0.8%, (0.4–1.6)          |
| Oro-tracheal intubation                      | 4                | 0.4%, (0.1–1.0)          |
| Death                                        | 109              | 10.1%, (8.4–12.1)         |
| Other                                        | 42               | 3.9%, (2.8–5.3)          |

**Causes of 15-day mortality**

- Death caused exclusively by PE: 19 episodes (1.8%, (1.1–2.8))
- PE-related death, mixed: 17 episodes (1.6%, (0.9–2.5))
- Fatal bleeding: 6 episodes (0.6%, (0.2–1.2))
- Cancer: 45 episodes (4.2%, (3.1–5.6))
- Infection: 8 episodes (0.7%, (0.3–1.5))
- Arterial thrombosis: 0 episodes (0)
- Other/unknown: 14 episodes (1.3%, (0.7–2.2))

**15-day venous rethrombosis**: 8 episodes (0.7%, (0.3–1.5))

**All-cause 30-day mortality**: 153 episodes (14.2%, (12.2–16.5))

**30-day serious bleeding**: 59 episodes (5.5%, (4.2–7.0))

**30-day venous rethrombosis**: 12 episodes (1.1%, (0.6–2.0))

Abbreviations: BP = blood pressure; CI = confidence interval; ICU = intensive care unit; PE = pulmonary embolism.

<sup>a</sup> These events are not mutually exclusive.

---

**Figure 1. Study flow diagram.** CT, computed tomography; PE, pulmonary embolism.
ECOG-PS, Eastern Cooperative Group Performance Status scale; PE, pulmonary embolism; SaO_2, arterial oxygen saturation.

Patients are classified as low or high risk according to whether they exhibit none of these characteristics or at least one of them. CDR, clinical decision rule; categories are significantly different (P-test reveals that the survival functions factored by prognostic cluster in the first hours following diagnosis of PE. The log-rank seen to be evenly distributed throughout the 15 days and do not raise progressively, in patients of low, intermediate, and high risk: P_o.3, 6.1, and 17.1%; P_o.717–0.840).

The EPIPHANY index is applicable across the entire spectrum of PE severity, including both incidental and symptomatic events (Wicki et al, 2000; Aujesky et al, 2006; Uresandi et al, 2007; Jiménez et al, 2010; Kline et al, 2012; Den Exter et al, 2013). The model is a validation and extension of the CDR proposed in several clinical trials with the aim of pragmatically selecting low-risk patients eligible for outpatient care (Siragusa et al, 2005; Zondag et al, 2011; Font et al, 2014; Weeda et al, 2016). These decision-making rules are based on the combination of altered vital signs (e.g., hypotension, hypoxaemia, tachycardia, etc.) and factors that point toward a high risk of bleeding or other contraindications to receiving treatment in the home. Moreover, the EPIPHANY rule incorporates another five covariates that include discriminatory characteristics typical in cancer patients that are easily accessible at patients’ bedside, such as ECOG-PS, evaluation of tumour response prior to PE using RECIST 1.1 criteria, previous primary tumour resection, oxygen saturation, and the presence or absence of PE-specific symptoms. All these variables have been widely used in various contexts to predict clinical outcome and there is good reason to think that they are also important in PE (Wicki et al, 2000; Aujesky et al, 2006; corrected value of the area under the ROC curve is 0.779 (95% CI, 0.717–0.840).

Outcomes according to risk groups are reported in Table 3. The risk of serious complications within 15 days increases with the group: 1.6, 9.4, 30.6%; P<0.0001. The risk of 15-day mortality also raises progressively, in patients of low, intermediate, and high risk: 0.3, 6.1, and 17.1%; P<0.0001. It is worth noting that high–intermediate-risk patients had increased risk of complications, with OR of 17.2 (95% CI, 7.7–40.3), P<0.0001, and death OR of 49.5 (95% CI, 6.8–356.9), P<0.0001.

Figure 3 illustrates the cumulative hazard function. Events are seen to be evenly distributed throughout the 15 days and do not cluster in the first hours following diagnosis of PE. The log-rank test reveals that the survival functions factored by prognostic categories are significantly different (P<0.0001).

**DISCUSSION**

This study reports on the development of a decision tree model to stratify any cancer patient with PE according to the risk of serious complications within 15 days. Unlike other prognostic tools, the EPIPHANY index is applicable across the entire spectrum of PE severity, including both incidental and symptomatic events (Wicki et al, 2000; Aujesky et al, 2006; Uresandi et al, 2007; Jiménez et al, 2010; Kline et al, 2012; Den Exter et al, 2013). The model is a validation and extension of the CDR proposed in several clinical trials with the aim of pragmatically selecting low-risk patients eligible for outpatient care (Siragusa et al, 2005; Zondag et al, 2011; Font et al, 2014; Weeda et al, 2016). These decision-making rules are based on the combination of altered vital signs (e.g., hypotension, hypoxaemia, tachycardia, etc.) and factors that point toward a high risk of bleeding or other contraindications to receiving treatment in the home. Moreover, the EPIPHANY rule incorporates another five covariates that include discriminatory characteristics typical in cancer patients that are easily accessible at patients’ bedside, such as ECOG-PS, evaluation of tumour response prior to PE using RECIST 1.1 criteria, previous primary tumour resection, oxygen saturation, and the presence or absence of PE-specific symptoms. All these variables have been widely used in various contexts to predict clinical outcome and there is good reason to think that they are also important in PE (Wicki et al, 2000; Aujesky et al, 2006;
Patients having any risk factor and poor functional status have a thrombotic risk (Jiménez et al, 1982). In general, functional worsening points to worse prognosis than those with a good functional status, particularly when PE diagnosis is incidental. On the other hand, we have detected a new variable that should be incorporated into the CDR: evaluation of tumour response prior to PE based on RECIST radiological criteria, which determines short-term prognosis following PE. Tumours in progression or those at risk for progression because response could not be assessed are at higher risk for complications than those with controlled disease or with no evidence of disease, even in the absence of other prognostic factors. In fact, resection of the primary tumour appears to be the only thing that protects individuals with tumours in progression and no other risk factor facing a complicated clinical course. This variable likely improves prognosis as a consequence of decreasing local complications, such as serious bleeding (Lee et al, 2009). In fact, in this series, bleeding was located in the primary tumour in 43% of the cases, which rose to 50% in subjects who died due to haemorrhage.

Remarkably, some patients diagnosed incidentally displayed PE-specific symptoms upon meticulous anamnesis, as reported by other authors (O’Connell et al, 2006). Therefore, the absence of PE-specific symptoms does not correspond exactly with incidental PE.

The decision tree model classification method used after the Exhaustive CHAID procedure is one of the differences that distinguish the EPIPHANY index from other models (Wicki et al, 2000; Aujesky et al, 2006; Uresandi et al, 2007; Jiménez et al, 2010; Kline et al, 2012; Den Exter et al, 2013). The ECOG-PS has been largely acknowledged in oncology to predict toxicity and adverse clinical outcomes in various contexts (Oken et al, 1982). In general, functional worsening points to severe underlying pathology, poorer physiological reserve, and decreased mobility, thereby making patients more prone to thrombotic risk (Jiménez et al, 2010; Den Exter et al, 2013). Patients having any risk factor and poor functional status have a worse prognosis than those with a good functional status, particularly when PE diagnosis is incidental. On the other hand, we have detected a new variable that should be incorporated into the CDR: evaluation of tumour response prior to PE based on RECIST radiological criteria, which determines short-term prognosis following PE. Tumours in progression or those at risk for progression because response could not be assessed are at higher risk for complications than those with controlled disease or with no evidence of disease, even in the absence of other prognostic factors. In fact, resection of the primary tumour appears to be the only thing that protects individuals with tumours in progression and no other risk factor facing a complicated clinical course. This variable likely improves prognosis as a consequence of decreasing local complications, such as serious bleeding (Lee et al, 2009). In fact, in this series, bleeding was located in the primary tumour in 43% of the cases, which rose to 50% in subjects who died due to haemorrhage.

Table 3. Outcomes in each risk group (n = 1075)

| Category                          | All patients, n = 1075 | Low risk: nodes 6, 12 | Intermediate risk: nodes 10, 11 | High risk: nodes 7, 8, 9 | P-value |
|-----------------------------------|------------------------|-----------------------|---------------------------------|-------------------------|---------|
| 15-day complications              |                        |                       |                                 |                         |         |
| Overall                           | 208 (19.3%)            | 6 (2.0%)              | 20 (9.4%)                       | 182 (32.7%)             | <0.0001 |
| Hypotension                       | 50 (4.7%)              | 0                     | 0                               | 50 (9.0%)               | <0.0001 |
| Acute respiratory failure         | 86 (8.0%)              | 1 (0.3%)              | 1 (0.5%)                        | 84 (15.1%)              | <0.0001 |
| Fibrinolysis                      | 7 (0.7%)               | 0                     | 0                               | 7 (1.3%)                | 0.019   |
| Major bleeding                    | 54 (5.0%)              | 3 (1.0%)              | 5 (2.3%)                        | 46 (8.3%)               | <0.0001 |
| Right-side heart failure          | 22 (2.0%)              | 0                     | 2 (0.9%)                        | 20 (3.6%)               | <0.0001 |
| Acute kidney failure              | 17 (1.6%)              | 0                     | 1 (0.5%)                        | 16 (2.9%)               | 0.001   |
| ICU admission                     | 21 (2.0%)              | 0                     | 0                               | 21 (3.8%)               | <0.0001 |
| Cardiopulmonary resuscitation     | 4 (0.4%)               | 0                     | 0                               | 4 (0.7%)                | 0.075   |
| Non-invasive mechanical ventilation| 9 (0.8%)               | 0                     | 0                               | 9 (1.6%)                | 0.008   |
| Oro-tracheal intubation           | 4 (0.4%)               | 0                     | 0                               | 4 (0.7%)                | 0.076   |
| Demise                            | 109 (10.1%)            | 1 (0.3%)              | 13 (6.1%)                       | 95 (17.1%)              | <0.0001 |
| Other serious complication        | 42 (3.9%)              | 1 (0.3%)              | 2 (0.9%)                        | 39 (7.0%)               | <0.0001 |
| Day first complication appeared, median (range) | 3 (0–15) | 13 (5–15) | 7 (0–15) | 3 (0–15) | 0.020   |
| Cause of death at 15 days         |                        |                       |                                 |                         |         |
| PE                                | 19 (1.8%)              | 0                     | 2 (0.9%)                        | 17 (3.1%)               | 0.001   |
| Mixed cause                       | 17 (1.6%)              | 0                     | 0                               | 17 (3.1%)               | <0.0001 |
| Bleeding                          | 6 (0.6%)               | 0                     | 1 (0.5%)                        | 5 (0.9%)                | 0.088   |
| Disease progression               | 45 (4.2%)              | 1 (0.3%)              | 10 (4.7%)                       | 34 (6.1%)               | <0.0001 |
| Sepsis                            | 8 (0.7%)               | 0                     | 0                               | 8 (1.4%)                | 0.012   |
| Arterial thrombosis               | 0                      | 0                     | 0                               | 0                       | —       |
| Other reason                      | 14 (1.3%)              | 0                     | 0                               | 14 (2.5%)               | 0.001   |
| Venous rethrombosis within 15 days| 8 (0.7%)               | 3 (1.0%)              | 0                               | 5 (0.9%)                | 0.959   |
| Recurrence of PE at 15 days       | 43 (4.0%)              | 16 (5.2%)             | 13 (6.1%)                       | 14 (2.5%)               | 0.030   |
| Arterial rethrombosis within 15 days| 25 (2.3%)         | 2 (0.7%)              | 8 (3.8%)                        | 15 (2.7%)               | 0.095   |
| 30-day mortality                  | 153 (14.2%)            | 4 (1.3%)              | 19 (8.9%)                       | 130 (23.3%)             | <0.0001 |

Abbreviations: ICU = intensive care unit; PE = pulmonary embolism. Percentages refer to columns. It was possible to record more than one serious complication for the same patient; causes of death could overlap. The P-value was calculated using the Linear-by-Linear association test, except for the day the complication appeared (independent samples median test).
and are better suited to elaborating decision-making algorithms that follow the same structure (Yohannes and Hoddinott, 1999; Lewis, 2000). Thus, in the real world, decisions in subjects with PE are not generally made on the basis of the small additive contributions of several variables, but on the presence or absence of strong dichotomous predictors such as cardiogenic shock, acute respiratory failure, hypotension, etc. (Wicki et al., 2000; Aujesky et al., 2006; Uresandi et al., 2007; Jiménez et al., 2010; Kline et al., 2012; Den Exter et al., 2013). The presence of a single one of these variables indicates high risk and is fundamental in the clinical decision to intensify therapy, regardless of the contribution of the remaining covariates of a logistic regression model. Decision trees are also useful in situations having non-linear probable effects for some variables, as is assumed in a sample of patients diagnosed using different methods (CT-angiography scans vs conventional CT) and having dissimilar clinical characteristics, depending on if they are incidental or symptomatic events. Insofar as the previously developed scales are concerned (RIETE, POMPE-C, PESI, etc.) (Aujesky et al., 2006; Jiménez et al., 2010; Kline et al., 2012; Den Exter et al., 2013), we do not know for sure if they can complement these criteria, although a preliminary analysis performed by our group suggests that their use is not likely to be necessary after applying a clinical classification rule (Carmona-Bayonas et al., 2016).

Another striking difference between the EPIPHANY index and the afore-mentioned methods is that we propose beginning to use the probability of serious complications within 15 days as the primary end point and not all-cause 30-day mortality, which is the probability of all-cause 30-day mortality will not necessarily affect decision-making regarding ambulatory treatment in some subgroups, as the cause of death is rarely the PE itself, and is often due to cancer progression (Den Exter et al., 2013; Carmona-Bayonas et al., 2016). In fact, patients on palliative care for advanced disease are those in whom it is even more important to prevent unnecessary hospitalisation at the end of their lives. The use of the all-cause 30-day mortality end point also entails the issue of proposing intensification of PE management (e.g., with fibrinolysis) in subjects at greater risk of early mortality due to cancer; who are precisely the ones who are less likely to benefit. For instance, when we examine the causes of death 15, 30, and 90 days after PE, cancer is the cause of death in 35%, 54%, and 65%, respectively. Although determining the cause of death in absence of autopsies has clear limitations, the same data appear in the RIETE registry, in which ~50% of the deaths resulted from the cancer itself and not PE (Den Exter et al., 2013).

This study has certain limitations that must be taken into account. First of all, it is a fundamentally retrospective registry of medical history data, with the intrinsic limitations in precision this entails. Nevertheless, most of the events contemplated are solid and are faithfully recorded in the histories (blood pressure, oxygen saturation, documentation of tumour response, ECOG-PS, exitus, etc.). Second, PE is a highly polymorphic pathology and more external validations are needed by other groups, being cognizant that these models offer a general overview of the main risk factors. However, some models have other particular factors with a definitive impact on prognosis. Third, decision tree modelling can be weak, unstable predictors in certain contexts. Thus, random forest models that incorporate the prediction of multiple, individual decision trees may perform better, albeit they are also more complicated to interpret and use in daily practice (Breiman, 2001). It is also doubtful that they can achieve a better definition of ‘low risk’. Finally, the assimilation and integration of radiological variables and/or biomarkers (e.g., troponin, pulmonary artery obstruction indices, right ventricular dilatation, etc.) would call for more in-depth studies.

In short, we have elaborated a decision tree to predict serious complications in cancer patients with PE that enables patients to be classified into groups of high, intermediate, and low risk for complications. This model validates and refines the classification rules previously used by other authors; it is based on variables that are easy to obtain; it’s easy to use, and can have potential implications for clinical management.

ACKNOWLEDGEMENTS

We thank Priscilla Chase Duran for editing the text. This project was funded in part by a restricted educational grant from Leo Pharma Spain and by support from the Asociación de Investigación de la Enfermedad Tromboembólica de la Región de Murcia (the Region of Murcia’s Association of Thromboembolic Disease Research).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Aujesky D, Obrosky DS, Stone RA, Auble TE, Perrier A, Cornuz J, Roy P-M, Fine MJ (2006) A prediction rule to identify low-risk patients with pulmonary embolism. Arch Intern Med 166: 169–175.

Biggs D, De Ville B, Suen E (1991) A method of choosing multivariate partitions for classification and decision trees. J Appl Stat 18: 49–62.

Breiman L (2001) Random forests. Mach Learn 45: 5–32.

Carmona-Bayonas A, Font C, Fonseca PJ, Fenoy F, Otero R, Beato C, Plasencia J, Biosca M, Sánchez M, Benegas M, Calvo-Temprano D, Varona D, Faez L, Vicente MA, de la Haba I, Antonio M, Madridano O, Ramchandani A, Castañón E, Marchena PJ, Martínez MJ, Martín M, Marin G, de la Peña FA, Vicente V (2016) On the necessity of new decision-making methods for cancer-associated, symptomatic, pulmonary embolism. Thromb Res 143: 76–85.

Den Exter PLPL, Gómez V, Jiménez D, Trujillo-Santos J, Muriel A, Huisman MV, Monreal M, Registro Informatizado de la Enfermedad Tromboembólica (RIETE) InvestigadoresGómez V, Jiménez D (2013) A clinical prognostic model for the identification of low-risk patients with acute symptomatic pulmonary embolism and active cancer. Chest 143: 138–145.

Dentali F, Ageno W, Becattini C, Galli L, Gianni M, Riva N, Imberti D, Squizzato A, Venco A, Agnelli G (2010) Prevalence and clinical history of incidental, asymptomatic pulmonary embolism: a meta-analysis. Thromb Res 125: 518–522.

Eisenhaauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd D, Kaplan R, Lacombe D, Verweij J (2009) New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 45: 228–247.

Font C, Carmona-Bayonas A, Beato C, Reig Ò, Sáez A, Jiménez-Fonseca P, Plasencia JM, Calvo-Temprano D, Sanchez M, Benegas M (2016) Clinical features and short-term outcomes of cancer patients with suspected and unsuspected pulmonary embolism: the EPIPHANY study. Eur Respir J 49: 1600282.

Font C, Carmona-Bayonas A, Fernández-Martínez A, Beato C, Vargas A, Gascon P, Otero R (2014) Outpatient management of pulmonary
Plasencia-Martínez JM, Carmona-Bayonas A, Calvo-Temprano D, Qanadli SD, El HM, Vieillard-Baron A (2001) New CT index to quantify pulmonary angiography indices in patients with acute symptomatic pulmonary embolism. Arch Intern Med 170: 1383–1389.

Kline IA, Roy PM, Than MP (2012) Derivation and validation of a multivariate model to predict mortality from pulmonary embolism with cancer: the POMPE-C tool. Thromb Res 129: e194–e199.

Lee JA, Lim DH, Park W, Ahn YC, Huh SJ (2009) Radiation therapy for arterial obstruction in pulmonary embolism: comparison with clinical correlates and relevance. J Clin Oncol 27: 749–755.

Lewis RJ (2000) An introduction to classification and regression tree (CART) analysis. In Annual Meeting of the Society for Academic Emergency Medicine; San Francisco, CA, USA, pp 1-14.

Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, Carbone PP (1982) Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol 5: 649–655.

O’Connell CL, Boswell WD, Duddalwar V, Caton A, Mark LS, Vigen C, Liebman HA (2006) Unsuspected pulmonary emboli in cancer patients: clinical correlates and relevance. J Clin Oncol 24: 4928–4932.

Plasencia-Martínez JM, Carmona-Bayonas A, Calvo-Temprano D, Jiménez-Fonseca P, Fenoy F, Benegas M, Sánchez M, Font C, Varona D, de la Haza DM (2016) Prognostic value of computed tomography pulmonary angiography indices in patients with cancer-related pulmonary embolism: data from a multicenter cohort study. Eur J Radiol 87: 66–75.

Qanadli SD, EI HM, Vieillard-Baron A (2001) New CT index to quantify arterial obstruction in pulmonary embolism: comparison with angiographic index and echocardiography. AIR Am J Roentgenol 176: 1415–1420.

Ranpuria V, Hapani S, Wu S (2011) Treatment-related mortality with bevacizumab in cancer patients: a meta-analysis. JAMA 305: 487–494.

Siragusa S, Arcara C, Malato A, Anastasio R, Valerio MR, Fulfarò F, Lo Coco I, Grimaudo S, Bajardi G, Abbadesa V, Gebbia N (2005) Home therapy for deep vein thrombosis and pulmonary embolism in cancer patients. Ann Oncol 16(Suppl 4): iv136–iv139.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)