View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

VALUE IN HEALTH 16 (2013) 1111–1122

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

ORIGINAL RESEARCH
*Economic Evaluation*

# An Integrated Approach to Evaluating Alternative Risk Prediction Strategies: A Case Study Comparing Alternative Approaches for Preventing Invasive Fungal Disease

*Z. Sadique, PhD[1,*], R. Grieve, PhD[1], D.A. Harrison, PhD[2], M. Jit, PhD[3,4], E. Allen, PhD[5], K.M. Rowan, DPhil[2], FIRE Study Investigators*

[1]*Department of Health Services Research & Policy, London School of Hygiene and Tropical Medicine, London, UK;* [2]*Intensive Care National Audit & Research Centre, London, UK;* [3]*Health Protection Agency, London, UK;* [4]*Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK;* [5]*Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK*

ABSTRACT

**Objectives:** This article proposes an integrated approach to the development, validation, and evaluation of new risk prediction models illustrated with the Fungal Infection Risk Evaluation study, which developed risk models to identify non-neutropenic, critically ill adult patients at high risk of invasive fungal disease (IFD). **Methods:** Our decision-analytical model compared alternative strategies for preventing IFD at up to three clinical decision time points (critical care admission, after 24 hours, and end of day 3), followed with antifungal prophylaxis for those judged "high" risk versus "no formal risk assessment." We developed prognostic models to predict the risk of IFD before critical care unit discharge, with data from 35,455 admissions to 70 UK adult, critical care units, and validated the models externally. The decision model was populated with positive predictive values and negative predictive values from the best-fitting risk models. We projected lifetime cost-effectiveness and expected value of partial perfect information for groups of parameters. **Results:** The risk prediction models performed well in internal and external validation. Risk assessment and prophylaxis at the end of day 3 was the most cost-effective strategy at the 2% and 1% risk threshold. Risk assessment at each time point was the most cost-effective strategy at a 0.5% risk threshold. Expected values of partial perfect information were high for positive predictive values or negative predictive values (£11 million–£13 million) and quality-adjusted life-years (£11 million). **Conclusions:** It is cost-effective to formally assess the risk of IFD for non-neutropenic, critically ill adult patients. This integrated approach to developing and evaluating risk models is useful for informing clinical practice and future research investment.

*Keywords:* cost-effectiveness analysis, critical care, invasive fungal disease, risk prediction, value of information analysis.

Copyright © 2013, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

## Introduction

Risk prediction models have great potential to support clinical decisions and the development of clinical guidelines [1–5]. For example, the decision to initiate statin therapy for the primary prevention of cardiovascular disease may be informed by risk equations from the Framingham study [6]. Treatment choice for patients with breast cancer can be guided by estimates of the long-term risk of cancer recurrence or death, for example, from the Nottingham prognostic index [7]. Clinical decision making in critical care units may be informed by estimates of the predicted risk of death, based, for instance, on the acute physiology and chronic health evaluation score [8,9]. In many circumstances, however, it is unclear whether using risk prediction approaches to initiate prevention and treatment strategies is cost-effective.

Risk prediction models can be used in cost-effectiveness analysis (CEA) to identify which patient subgroups are the most cost-effective to receive a particular treatment or prevention strategy [10–12]. For example, Grieve et al. [13] considered alternative Framingham equations to evaluate strategies for preventing cardiovascular disease, and Williams et al. [14] outlined the use of a prognostic model to select patients with breast cancer for systemic therapy. Longworth et al. [15] used published risk prognostic models to evaluate the cost-effectiveness of liver transplantation. None of these studies, however, evaluated whether a strategy of formal risk assessment with a prognostic model was cost-effective. Furthermore, previous CEAs have taken a published risk prediction model and assumed that it is valid for the decision context. The population characteristics in the decision context are often different from those of the population, on

* *Address correspondence to:* Z. Sadique, Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK.

E-mail: Zia.Sadique@lshtm.ac.uk.

1098-3015/$36.00 – see front matter Copyright © 2013, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.jval.2013.09.006

which the risk prediction model was developed, leading to unreliable extrapolations. To assess whether a risk prediction model is valid in the specific decision context requires a careful assessment of the statistical performance of the model for the relevant population and time points. In particular, it is important to consider whether the risk prediction model accurately predicts events not only for the original population on which it was developed (internal validation) but also for alternative populations of prime interest for the decision problem (external validation). Of greatest importance for decision making is the discrimination of the risk model. If a risk model has perfect discrimination, then there is a threshold risk that divides the patients into those who will versus those who will not experience an event, leading to the optimal treatment decision for every patient. In practice, perfect discrimination will not be achieved, but improved discrimination will lead to better decision making, reflected through the positive predictive value (PPV) and the negative predictive value (NPV) of the decision rule. In addition to developing risk models that are accurate, it is important to evaluate the relative cost-effectiveness of alternative risk prediction approaches.

There is growing appreciation of the need to evaluate risk prediction models, but only a few studies have done this, and only to a limited extent. Henriksson et al. [16] developed a risk prediction model and assessed the cost-effectiveness of prognostic biomarkers and risk scores to inform prioritization for coronary artery surgery. Rapsomaniki et al. [17] evaluated the net benefit from using a prognostic model, illustrated in the context of the prevention of cardiovascular disease. But none of the above studies has fully assessed the uncertainty in the decision problem by assessing the value of information to help support decision making. Value of information analysis provides an important framework for determining the expected payoff of conducting further research to resolve the parameter uncertainties that pervade the cost-effectiveness estimates [18]. We propose an integrated approach to considering risk prediction models in decision making. The integrated approach considers developing, validating, and evaluating the cost-effectiveness of a risk prediction approach for the relevant decision context, for example, according to the specific population and time point of interest. This integrated approach also requires that the ensuing decision uncertainty be fully recognized by providing an assessment of the priorities for further research.

The integrated approach is illustrated with the Fungal Infection Risk Evaluation (FIRE) study, which developed prognostic models to identify non-neutropenic, critically ill adult patients at high risk of invasive fungal disease (IFD). For critically ill patients, IFD is associated with increased morbidity, mortality, and cost [19–21]. Randomized controlled trials have reported that antifungal prophylaxis with either fluconazole or ketoconazole reduces the subsequent risk of IFD and mortality [22]. These randomized controlled trials were conducted in high-risk patients, and concerns about the costs of prophylaxis and possible drug resistance have discouraged the widespread adoption of antifungal prophylaxis. In the United Kingdom, risk models are not routinely used to identify those non-neutropenic, critically ill adult patients who are at high risk of IFD. Antifungal prophylaxis is prescribed only on an ad-hoc basis for those patients who, according to clinical judgment, are at very high risk of IFD. In the FIRE study, only 1% of eligible patients received systemic antifungal therapy at admission to the critical care unit [23]. For the vast majority of patients admitted to critical care units who do not currently receive antifungal prophylaxis, it is unknown whether it is cost-effective to formally assess the risk of IFD at different clinical decision time points, and to initiate antifungal prophylaxis for those judged high risk.

The objective of this article was to illustrate an integrated approach to the development, validation, and CEA of risk prediction models through the FIRE case study. We use these risk prediction models to report the relative cost-effectiveness of alternative risk assessment and prophylaxis strategies for preventing IFD and assess the relative value of further research.

The article proceeds as follows. The next section outlines the decision, problem, and the CEA model, followed by a summary of risk model development and validation. Following this are sections on methods and results of CEA, scenario analysis, and value of information (VOI) analysis. In the final section, we discuss the approach taken and suggest a research agenda.

## Overview of the Decision Problem and the CEA Model

The CEA aimed to assess the relative cost-effectiveness of alternative strategies for assessing the risk of IFD and initiating antifungal prophylaxis in non-neutropenic, adult patients admitted to National Health Service critical care units in the United Kingdom. The CEA reported cost-effectiveness over the patients' lifetime and assessed costs from the National Health Service perspective. The alternative prevention strategies comprised "formal risk assessment" according to the predicted risk of IFD at up to three clinical decision time points (from herein termed "risk assessment"). These time points were at critical care unit admission, after 24 hours, and at the end of the third calendar day in the critical care unit (Table 1).

At any clinical decision time point, risk assessment was considered only for those patients who were still in the critical

**Table 1 – Alternative treatment strategies for non-neutropenic, critically ill adult patients.**

| Strategy | Decision node | | |
|---|---|---|---|
| | On admission | At end of 24 h | At end of day 3 |
| 1 | Do not assess risk | Do not assess risk | Do not assess risk |
| *Risk assessment at a single time point* | | | |
| 2 | Assess risk, Prophylaxis if risk > $P_T$ | Do not assess risk | Do not assess risk |
| 3 | Do not assess risk | Assess risk, Prophylaxis if risk > $P_T$ | Do not assess risk |
| 4 | Do not assess risk | Do not assess risk | Assess risk, Prophylaxis if risk > $P_T$ |
| *Risk assessment at multiple time points* | | | |
| 5 | Assess risk, Prophylaxis if risk > $P_T$ | Assess risk, Prophylaxis if risk > $P_T$ | Do not assess risk |
| 6 | Do not assess risk | Assess risk, Prophylaxis if risk > $P_T$ | Assess risk, Prophylaxis if risk > $P_T$ |
| 7 | Assess risk, Prophylaxis if risk > $P_T$ | Do not assess risk | Assess risk, Prophylaxis if risk > $P_T$ |
| 8 | Assess risk, Prophylaxis if risk > $P_T$ | Assess risk, Prophylaxis if risk > $P_T$ | Assess risk, Prophylaxis if risk > $P_T$ |

$P_T$, risk threshold.

care unit, not already receiving systemic antifungal therapy, and without IFD before that time point. Patients whose predicted risk was greater than a prespecified threshold ($P_T$) were designated "high risk." The risk thresholds defined a priori according to the literature and expert opinion were 0.5%, 1%, 2%, 5%, and 10%. It was assumed that antifungal prophylaxis was initiated for those newly defined as high risk at the particular clinical decision time point. The prophylaxis treatment regimen was assumed to follow current recommendations and was for 400 mg fluconazole per day for 10 days [22–25]. There is no specific guideline on the duration of antifungal prophylaxis in critical care. A systematic review [22] suggested that such prophylaxis is generally administered until discharge from critical care, but the duration varied across studies. Our study assumed that prophylaxis was administered for 10 days, the average length of stay in critical care for the study population. These strategies were contrasted to the current practice of "no formal risk assessment or prophylaxis" (Table 1) (from herein termed "no risk assessment").

A decision analytical model (Fig. 1) was developed to evaluate the alternative strategies for assessing the risk of IFD defined in Table 1. Alternative strategies at three clinical time points were compared with current practice, leading to eight alternative strategies under consideration. Each time point defined a clinical decision node, with two possibilities—either formal assessment of the risk of IFD was not undertaken (no risk assessment) or risk assessment was undertaken (risk assessment). If risk assessment was not undertaken, patients at each time point $i$ faced the risk of either having IFD in the critical care unit ($R_i$) or not having IFD ($1 - R_i$). From the "no IFD" health state, patients faced a baseline risk of all-cause death, and for patients predicted to develop IFD, an excess risk of death was applied (see the section on CEA of alternative risk assessment strategies).

Under the strategies in which risk assessment was undertaken, the proportion of patients ($P_i$) whose predicted risk of infection was higher than the risk threshold were judged high risk and assumed to receive prophylaxis. For these patients, the probability of developing IFD at any time during the stay on the critical care unit was estimated from time point–specific positive predictive value ($PPV_i$) and the relative risk of IFD following antifungal prophylaxis versus no prophylaxis [22]. The proportion of patients ($1 - P_i$) whose predicted risk of infection was lower than the risk threshold were judged "low risk" and assumed not to receive prophylaxis. For these patients, the probability of developing IFD was estimated as 1 minus the NPV ($1 - NPV_i$). The risk of death, conditional on the presence or absence of IFD, was assumed the same whether or not patients received antifungal prophylaxis.

The decision model required PPVs and NPVs for each clinical decision time point because the patients who did not receive antifungal prophylaxis may be reconsidered for risk assessment and prophylaxis at subsequent clinical decision time points.

## Risk Model Development and Validation

### Overview of the FIRE Study

The FIRE study collected data on 60,778 admissions to 96 adult, general critical care units in the United Kingdom between July 2009 and March 2011 [23]. For the development of the risk models and the CEA, the following exclusion criteria were applied: age younger than 18 years, readmissions, neutropenia, active hematological malignancy, admission following solid organ transplant, receipt of systemic antifungal or IFD identified before the clinical decision time point, and death or discharge from the critical care unit before the clinical decision time point (for the second and third time points).

IFD was defined as a blood culture or sample from a normally sterile site positive for yeast or mould cells in a microbiological or histopathological report. For risk model development and CEA, IFD was restricted to infection with *Candida* species, which accounts for 94% of the cases of IFD.

### Model Development and Validation

Before evaluating the relative cost-effectiveness of risk prediction models, it is important to validate the alternative models in the appropriate decision context [26–29]. The model development and validation are required to discriminate those patients who will versus those who will not experience an event according to a particular risk threshold, and so provide key parameters for the decision model (e.g., the PPV and the NPV). In the FIRE study, we performed internal validation, temporal validation, and external validation. The alternative risk prediction models are initially subject to internal validation. Here, the data set is split into two parts, and cross-validation is undertaken, whereby the model is developed on the first portion (often called a development data set) and predictive accuracy of the model is assessed on the second portion of the data (known as the validation data set). In temporal validation, the performance of the model is assessed on subsequent patients from the same context, for example, the same geographical location. The internal and temporal validation of risk models is necessary but insufficient; a true evaluation of the generalizability of the risk prediction model requires evaluation on data from external sources (external validation) [26,27].

In the FIRE study, risk models were developed to predict the risk of IFD at the three clinical decision time points required for the CEA. Here, we provide a brief summary; full details are provided elsewhere [23,30]. The selection of risk factors for IFD was informed by a systematic literature review [31]. The data set was divided into the following development and validation samples: 1) development sample—all admissions to a random sample of participating critical care units in England, Wales, and Northern Ireland, July 2009 to December 2010; 2) random validation sample—all admissions to the remaining units in England, Wales, and Northern Ireland; 3) temporal validation sample—all admissions to units in the development sample, January to March 2011; and 4) geographical validation sample—all admissions to units in Scotland (Table 2). A sample size of 40,000 was selected for model development on the basis of 20 events per variable assuming 20 candidate predictors and an event rate of 1%. This sample size provided 80% power to detect, as statistically significant ($P < 0.05$), a risk factor present in 10% of the population associated with a 50% increase in the risk of IFD. A further 20,000 patients were included for external validation on the basis of a split of development to validation data of two-thirds to one-third.

Logistic regression models with a backward stepwise approach were used to model the risk of subsequently developing IFD. The logistic regression method was chosen because this is the established statistical technique for developing predictive models for critical care [32]. There are concerns that stepwise logistic regression method can lead to biased and unstable results [33–35], and so we used bootstrapping to guard against overfitting [26–28,36], and in the CEA model used the estimates of the NPV and the PPV from the validation rather than development data sets. Robust standard errors from the Huber-White estimator [37,38] were used to account for clustering of patients within critical care units. All candidate variables identified from univariable analysis were included in a multivariable model, and the model was progressively simplified by using backwards stepwise selection. The final risk model at admission included the following variables: admission for presurgical preparation; surgery within up to 7 days before admission (elective/scheduled with no unexpected complications, elective/scheduled with unexpected complications, emergency/

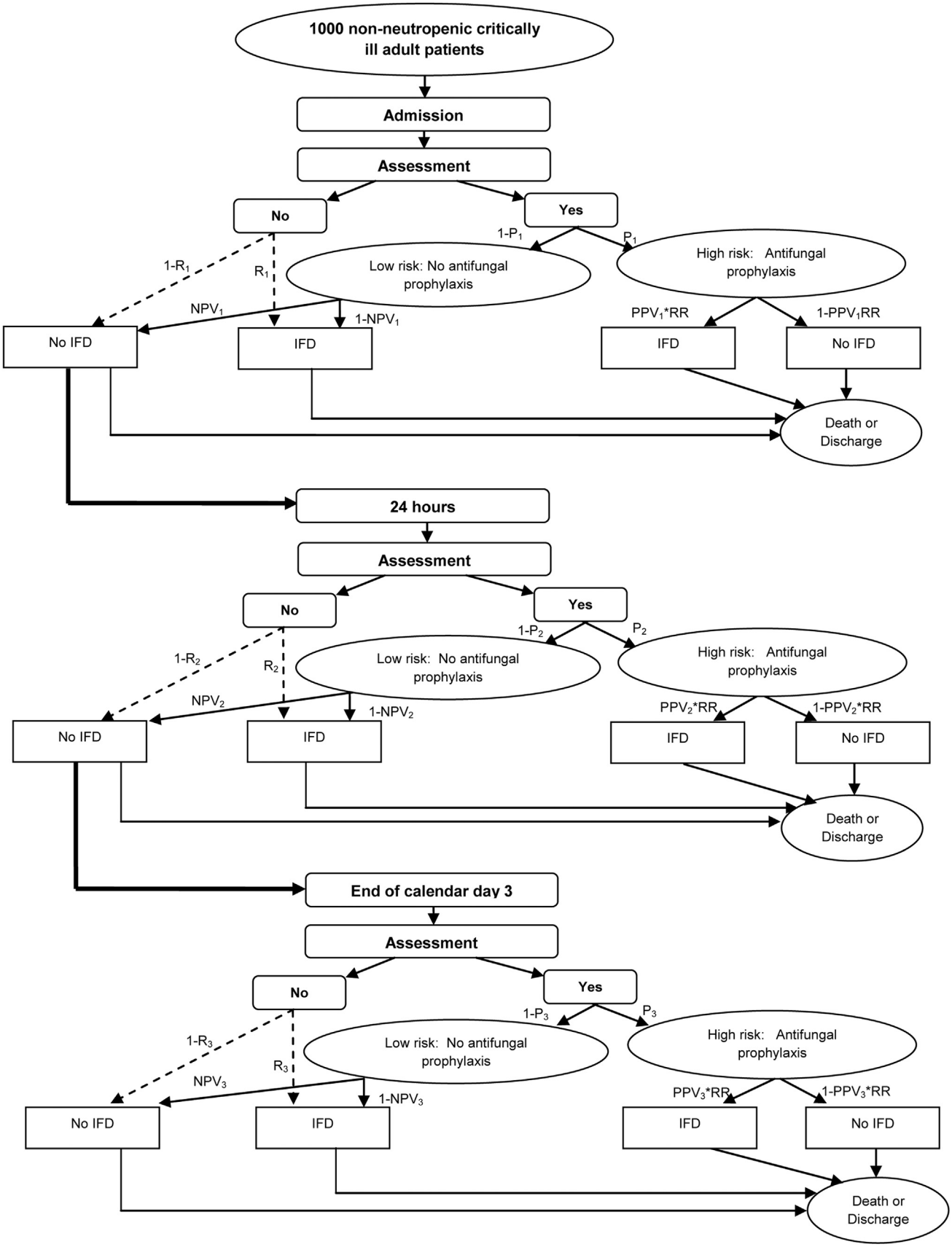


**Fig. 1 – Structure of the decision model comparing alternative strategies for assessing risk of IFD. IFD, invasive fungal disease; NPV, negative predictive value; PPV, positive predictive value, RR, relative risk.**

**Table 2 – Measures of risk prediction model performance in the development and validation samples.**

| | Development sample (n = 35,455) | | Validation sample | | | |
|---|---|---|---|---|---|---|
| | Original | Optimism adjusted | Random validation sample (n = 4,186) | Temporal validation sample (n = 9,866) | Geographical validation sample (n = 4,782) | Combined validation sample (n = 18,834) |
| At admission | | | | | | |
| c index* | 0.705 | 0.688 | 0.721 | 0.650 | 0.640 | 0.655 |
| Brier's score† | 0.0040 | 0.0041 | 0.0026 | 0.0043 | 0.0040 | 0.0038 |
| 24 h | | | | | | |
| c index | 0.824 | 0.810 | 0.840 | 0.759 | 0.650 | 0.732 |
| Brier's score | 0.0038 | 0.0038 | 0.0019 | 0.0042 | 0.0044 | 0.0037 |
| End of day 3 | | | | | | |
| c index | 0.835 | 0.825 | 0.803 | 0.720 | 0.661 | 0.709 |
| Brier's score | 0.0050 | 0.0050 | 0.0026 | 0.0049 | 0.0048 | 0.0043 |

* The c index is equivalent to the area under the receiver operating characteristic curve. A c index value of 0.5 indicates that the model is no better at predicting the outcome than random chance, whereas a value of 1 suggests perfect discrimination.

† Brier's score is the mean square error of the probability forecast over the validation sample and ranges between 0 (perfect predictions) and 0.25 (constant prediction of 0.5 for all patients).

urgent, no surgery); pancreatitis; number of catheters in central veins; number of drains; enteral feeding tube; and number of samples positive for fungal colonization. A full discussion of variables included in risk prediction models at other clinical decision time points is available elsewhere [23].

At each stage, the model was fitted in the development sample and the performance of the model was assessed. Model discrimination was assessed with the c index [34], and the overall fit was assessed by using the Brier's score [39]. Bootstrapping was used to internally validate the final model at each clinical decision time point and to estimate optimism-adjusted measures of discrimination and overall fit [40].

The final selected model at each clinical decision time point was evaluated in the three external validation samples by using the same performance measures as in the development sample. Analyses were performed by using Stata Version 10.1 (StataCorp LP, College Station, TX).

### *Results of Development and Validation of Risk Models*

In total, 144 admissions (0.4%) in the development sample had IFD. The risk model at admission had fair discrimination (c index 0.705). Discrimination improved at 24 hours (c index 0.824), and this was maintained at the end of calendar day 3 (c index 0.835). Despite the large sample size, the low incidence of IFD in the FIRE study made robust statistical modeling difficult. It is possible that the risk models were overfitted, and this may contribute to the drop in model performance when assessed in the validation samples (Table 2).

## CEA of Alternative Risk Assessment Strategies

### *Estimation of Parameters for CEA Model*

A decision analytical model was developed to consider current clinical practice in the United Kingdom, which is "no formal risk assessment and no prophylaxis," versus seven comparator strategies involving formal risk assessment and prophylaxis at various time points. Because for this decision problem it was not necessary to consider recurrent probabilities of events over time, we used a simple decision tree rather than a broader structure such as a Markov model. The CEA model included a hypothetical cohort of 1000 homogeneous cases with characteristics defined by the patients who met the FIRE study inclusion criteria (see Table 1 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2013.09.006). The decision model required the PPV and the NPV for each strategy and for each risk threshold. The low incidence of IFD in the FIRE study meant that at the higher prespecified risk thresholds of 5% and 10%, no patients designated at high risk were predicted to have IFD, and hence at these higher thresholds the PPV and NPV parameters could not be estimated for the risk assessment strategies. Hence, the risk thresholds considered in the CEA were 0.5%, 1%, and 2%.

To avoid concerns about overfitting from using the FIRE study development data set, PPVs and NPVs were estimated from the validation sample only (n = 18,834). Table 3 presents PPVs and NPVs for each strategy and risk threshold. The PPVs were low for all strategies; even at the 2% risk threshold and with prophylaxis at each clinical decision time point, the PPVs remained below 2%. By contrast, the NPVs all exceeded 99%.

The baseline risk of infection (R1, R2, R3) and the baseline risk of death were estimated for the three clinical decision time periods from the combined FIRE study development and validation samples (n = 54,289) (Table 4). For patients with IFD, the excess risk of death was estimated from the combined FIRE data set. The relative risk of IFD after prophylaxis was taken from the Cochrane systematic review by Playford et al. [22]. The systematic review reported similar relative risks across different levels of baseline risk. Hence, we applied the same relative risk of IFD after prophylaxis for all clinical decision time points and all risk thresholds.

### *Costs, Life-Years, and Quality-Adjusted Life-Years*

The cost of risk assessment assumed that, according to discussion with three clinical experts in critical care, 10 minutes of nursing time was required per patient, giving a cost of £8.67 [41]. Prophylaxis costs were calculated by assuming a standard regimen of fluconazole 400 mg for 10 days with unit costs taken from the British National Formulary [42]. The unit cost of prophylaxis recognizes that according to the British National Formulary, nonproprietary fluconazole intravenous infusion was available from September 2011. Lengths of stay in the critical care unit and hospital were costed with corresponding unit cost per bed day

**Table 3 – The PPV and the NPV according to strategy and risk threshold.**

| Strategy | Clinical decision time point | The PPV by risk threshold | | | The NPV by risk threshold | | |
|---|---|---|---|---|---|---|---|
| | | 0.5% | 1.0% | 2.0% | 0.5% | 1.0% | 2.0% |
| 1 | No risk assessment | – | – | – | – | – | – |
| 2 | On admission | 0.85% | 1.94% | 1.32% | 99.95% | 99.94% | 99.92% |
| 3 | At end of 24 ho | 0.88% | 1.60% | 1.35% | 99.92% | 99.93% | 99.92% |
| 4 | At end of day 3 | 0.95% | 1.21% | 1.26% | 99.79% | 99.73% | 99.65% |
| 5 | On admission | 0.85% | 1.94% | 1.32% | 99.95% | 99.94% | 99.92% |
| | At end of 24 h | 0.70% | 1.31% | 1.71% | 99.92% | 99.93% | 99.92% |
| 6 | At end of 24 h | 0.88% | 1.60% | 1.35% | 99.92% | 99.93% | 99.92% |
| | At end of day 3 | 0.98% | 0.79% | 1.62% | 99.78% | 99.72% | 99.66% |
| 7 | On admission | 0.85% | 1.94% | 1.32% | 99.95% | 99.94% | 99.92% |
| | At end of day 3 | 0.70% | 0.99% | 1.38% | 99.78% | 99.73% | 99.64% |
| 8 | On admission | 0.85% | 1.94% | 1.32% | 99.95% | 99.94% | 99.92% |
| | At end of 24 h | 0.70% | 1.31% | 1.71% | 99.92% | 99.93% | 99.92% |
| | At end of day 3 | 0.57% | 0.85% | 1.70% | 99.77% | 99.73% | 99.66% |

*Note*: Data from Harrison D, Muskett H, Harvey S, et al. [23]. FIRE study full validation sample (n = 18,805) (excludes 29 patients with missing values).
FIRE, Fungal Infection Risk Evaluation; NPV, negative predictive value; PPV, positive predictive value.

**Table 4 – Input parameters for CEA model evaluating alternative risk assessment strategies for preventing IFD.**

| Parameter | Time period | Point estimate | Distribution* | Source |
|---|---|---|---|---|
| Baseline risk of IFD | Within 24 h† | 0.082% | Beta (50; 60,728) | Combined FIRE data set‡ |
| | 24–48 h† | 0.087% | Beta (32; 36,911) | |
| | After 72 h† | 0.449% | Beta (102; 22,624) | |
| Probability of death, no IFD | Within 24 h | 4.57% | Beta (2,777; 57,999) | Combined FIRE data set |
| | 24–48 h | 5.82% | Beta (2,151; 34,790) | |
| | After 72 h | 8.78% | Beta (3,057; 31,737) | |
| Relative risk of death, IFD vs. no IFD | During critical care unit | 2.14 | lognormal (0.57 to 0.95) | Systematic review [22] |
| Relative risk of IFD, prophylaxis vs. no prophylaxis | During critical care unit | 0.46 | lognormal (−1.17 to −0.39) | Systematic review [22] |
| Cost of course of prophylaxis (£) | In critical care unit | 77.80 | Gamma (0.98; 79.00) | British National Formulary [42] |
| Critical care unit LOS (d): no IFD | In critical care unit | 10.16 | Gamma (0.83; 12.18) | Combined FIRE data set |
| Critical care unit LOS (d): IFD | In critical care unit | 24.95 | Gamma (1.82; 13.74) | Combined FIRE data set |
| Hospital LOS (d): no IFD | After critical care unit | 22.72 | Gamma (0.57; 39.53) | Combined FIRE data set |
| Hospital LOS (d): IFD | After critical care unit | 36.60 | Gamma (1.07; 34.22) | Combined FIRE data set |
| Unit cost of critical care unit bed-day (£): no IFD | | 1,085 | Gamma (22.3; 48.63) | Reference cost by HRG [44] |
| Unit cost of critical care unit bed-day (£): IFD | | 1,351 | Gamma (31.06; 43.48) | Reference cost by HRG [44] |
| QALY§ | | 10.52 | Gamma (127.96; 0.08) | Published sources [45–47] |

CEA, cost-effectiveness analysis; CI, confidence interval; FIRE, Fungal Infection Risk Evaluation; HRG, Health Resource Group; IFD, invasive fungal disease; LOS, length of stay; QALY, quality-adjusted life-year.
* For beta distribution, the first parameter refers to the number of events (alpha) and the second one refers to N – alpha. For gamma distribution, the first parameter refers to mean$^2$/SE$^2$ and the second parameters refers to SE$^2$/mean, where SE represents standard error. For lognormal distribution, the values in parentheses represent 95% CI of mean.
† These risks are calculated at specific time points. Overall, baseline risk of infection at any time during the study is 0.4%.
‡ Combined FIRE data set includes both development and validation data sets. See Table 2 for details.
§ QALY is calculated at mean age of patients (i.e., 60 y).

from National Health Service reference costs and other published sources [43,44] (Table 4). All costs were adjusted to 2010–2011 price levels [41].

Life-years were calculated from the UK life tables by applying all-cause mortality after hospital discharge from previous studies and applied an excess mortality of 20% for up to 4 years [45–47]. The resultant life-years were combined with estimates of health-related quality of life (HRQOL) for the UK general population [48] to project lifetime quality-adjusted life-years (QALYs) for each patient. To recognize that the HRQOL of critical care unit survivors is lower than that of the general population, we assumed that the HRQOL was 80% that of the age-gender–matched general population [45]. Lifetime QALYs were calculated for patients aged 60 years at admission, approximately the mean age of patients included in the FIRE study. Future costs and outcomes were discounted at the rate of 3.5% recommended by the National Institute for Health and Care Excellence [4]. We used Microsoft Excel 2007 for the cost-effectiveness analysis.

### *Implementation*

Incremental costs, incremental QALYs, and incremental net benefits (INBs), at a threshold of £20,000 per QALY, were calculated as the differences in mean end points following each risk assessment strategy versus the "no risk assessment." A probabilistic sensitivity analysis was conducted by resampling the input parameters 5000 times from recommended probability distributions [49,50] (Table 4). Cost-effectiveness acceptability curves were calculated according to the proportion of replications for which each strategy was the most cost-effective, that is, had the maximum INBs across all eight strategies, at different levels of willingness to pay for a QALY gain (£0–£50,000 per QALY gained). The analyses were repeated for the risk thresholds of 0.5%, 1%, and 2%.

### *Base-Case Analysis CEA Results*

For the no risk assessment strategy, the mean total costs per patient were £16,772, with mean lifetime QALYs of 8.63 (see Table 2

**Table 5 – Incremental cost (£), QALYs, and INBs (£) for each prophylaxis strategy versus no risk assessment for alternative risk thresholds.**

| Strategy | Risk assessment thresholds and clinical decision time points | | Incremental cost mean | Incremental QALY* mean | INB† mean (95% credible intervals) |
|---|---|---|---|---|---|
| 1 | No risk assessment | | – | – | – |
| | Risk threshold of 0.5% | Time of risk assessment | | | |
| 2 | | On admission | 10 | 0.0004 | −2 (−60 to 65) |
| 3 | | At end of 24 h | 15 | 0.0000 | −16 (−90 to 58) |
| 4 | | At end of day 3 | −7 | 0.0017 | 41 (−81 to 162) |
| 5 | | On admission and at end of 24 h | 18 | 0.0005 | −7 (−92 to 78) |
| 6 | | At end of 24 h and at end of day 3 | −6 | 0.0016 | 38 (−89 to 165) |
| 7 | | On admission and at end of day 3 | −5 | 0.0020 | 46 (−81 to 172) |
| 8 | | On admission, at end of 24 h, and at end of day 3 | −4 | 0.0021 | 47 (−100 to 194) |
| | Risk threshold of 1% | Time of risk assessment | | | |
| 2 | | On admission | 9 | 0.0001 | −8 (−59 to 44) |
| 3 | | At end of 24 h | 12 | 0.0001 | −15 (−79 to 50) |
| 4 | | At end of day 3 | −7 | 0.0013 | 33 (−90 to 157) |
| 5 | | On admission and at end of 24 h | 19 | 0.0001 | −17 (−103 to 69) |
| 6 | | At end of 24 h and at end of day 3 | −5 | 0.0014 | 32 (−109 to 173) |
| 7 | | On admission and at end of day 3 | −3 | 0.0015 | 33 (−105 to 171) |
| 8 | | On admission, at end of 24 h, and at end of day 3 | 0 | 0.0016 | 32 (−132 to 195) |
| | Risk threshold of 2% | Time of risk assessment | | | |
| 2 | | On admission | 8 | 0.0001 | −7 (−22 to 8) |
| 3 | | At end of 24 h | 9 | 0.0000 | −8 (−29 to 12) |
| 4 | | At end of day 3 | −4 | 0.0008 | 20 (−23 to 64) |
| 5 | | On admission and at end of 24 h | 18 | 0.0001 | −16 (−42 to 9) |
| 6 | | At end of 24 h and at end of day 3 | 2 | 0.0009 | 16 (−32 to 64) |
| 7 | | On admission and at end of day 3 | 6 | 0.0008 | −10 (−36 to 57) |
| 8 | | On admission, at end of 24 h, and at end of day 3 | 10 | 0.0009 | 8 (−47 to 63) |

INB, incremental net benefit; QALY, quality-adjusted life-year.

* Incremental QALYs are rounded to four decimal places.

† Incremental costs and INBs are rounded to the nearest £. INBs are reported for the threshold of £20,000 per QALY.

in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2013.09.006). Compared with the no risk assessment strategy, each risk assessment strategy had positive but only small gains in lifetime QALYs. The incremental costs of the risk assessment strategies were negative for strategies with risk assessment at the end of calendar day 3, whether at single or at multiple time points. The INB at the risk threshold of 0.5% was highest when assessment and prophylaxis were administered at all time points. For the risk threshold of 1% to 2%, the highest INB was associated with risk assessment and prophylaxis at the end of calendar day 3 (Table 5). The most cost-effective risk assessment strategy differed according to the threshold risk of IFD. At a 1% or 2% risk threshold, risk assessment and prophylaxis at the end of calendar day 3 was the most cost-effective strategy at the recommended cost-effectiveness threshold of £20,000 per QALY gain. For the lower risk threshold (0.5%), the strategy with the highest probability of being cost-effective at £20,000 per QALY was to assess risk at each time point. At each risk threshold, there was considerable uncertainty surrounding the estimates of the relative cost-effectiveness of the alternative strategies, and at the £20,000 per QALY threshold, the probability that any particular strategy would be the most cost-effective did not exceed 30% (Fig. 2).

## Scenario Analysis

The main assumptions made in the base-case analysis were tested in the following scenario analyses:

a. 5 or 20 minutes of nursing time to undertake the risk assessment was considered.
b. Reduction of 75% in the unit cost of fluconazole. The rationale for this cost reduction was evidence from the systematic review that some randomized controlled trials have shown similar effectiveness with 100 mg rather than 400 mg of fluconazole per day [22]. Another rationale for considering a 75% reduction in the unit costs of fluconazole is that even if the dose is maintained at 400 mg, local discounts in the range of 50% to 70% may be available, and also fluconazole may be provided in tablet form once the patient is able to absorb the oral dose of fluconazole.
c. The duration of excess mortality was extended for up to 25 years [51].
d. The HRQOL of critical care unit survivors was assumed to be 70% (rather than 80%) that of the general population.
e. 5 or 14 days of prophylaxis duration as informed by systematic review [22] was considered.
f. Best-case and worst-case scenario as defined below were considered: Best-case scenario: prophylaxis administered for 5 days, 10 minutes of nursing time, 75% reduction in the prophylaxis unit cost, the HRQOL of survivors is 80% that of the general population, and excess mortality for 4 years.

Worst-case scenario: prophylaxis administered for 14 days, 20 minutes of nursing time, full cost of prophylaxis, the HRQOL of survivors is 70% that of the general population, and excess mortality for 25 years.

The scenario analysis reported that the base-case results were generally robust to the alternative assumptions considered (see Table 3 in Supplemental Materials found at http://dx.doi.org/10.1016/j.jval.2013.09.006). Across risk thresholds and risk assessment strategies, the scenarios that considered increased costs of assess ment and prophylaxis (e.g., higher nursing time, higher duration of prophylaxis) led to lower INB for the risk assessment strategies than for the base case, assuming lower costs of assessment and prophy laxis (e.g., lower nursing time, lower duration of prophylaxis) led to higher INB. Excess mortality for 25 years and decrement in "qual ity-of-life" weights shows a small effect on the INB. The general conclusion that the strategies that included risk assessment at the end of calendar day 3 were relatively cost-effective was robust to the alternative "best-case" and "worst-case" scenarios considered.

## VOI Analysis

There is always the possibility that the preventative strategy that appears the most cost-effective from current evidence would not be the optimal approach if perfect information was available. The expected costs of this decision uncertainty were quantified according to the expected value of perfect information (EVPI) [52]. The EVPI was calculated for the total population in the United Kingdom that would be anticipated to benefit from the strategies considered. We assumed that the eligible population of interest was 100,000 annual admissions to critical care units and that the relevant life cycle for the technology was 5 years [23]. The results of the ensuing VOI analyses are conditional on model structure and probability distributions for each parameter. The EVPI sets an upper bound on the return of resolving all the parameter uncertainties within the decision problem and referred to as an upper bound on the value of further research.

At a threshold of £20,000 per QALY, the total EVPI estimates ranged between £12 million (0.5% risk threshold) and £14 million (1% risk threshold). Hence, across all parameters in the decision model, the upper value of further research for the whole population of interest is high relative to the anticipated research costs. To establish where further research might be best targeted, we also reported the EVPI for parameters, or the expected value of partial perfect information [18]. The groups of parameters considered were baseline probability of IFD, mortality with and without IFD, PPV and NPV, relative risks of infection after prophylaxis, morbidity costs, and lifetime QALYs.

Figure 3 reports the expected value of partial perfect information estimates for each group of parameters by risk threshold. The VOI for each group of parameters is similar across the risk thresholds. The results also suggest that even after the FIRE study, given the large population of interest for this decision problem, further research to provide more precise estimates for parameters such as the PPV or the NPV (£11 million–£13 million) and QALYs (£11 million) is of high value.

## Discussion

This article presents an integrated approach to the development of prognostic models for clinical decision making. The integrated approach incorporates the development, validation, and evaluation of alternative risk prediction models within the same study. The approach is exemplified by the FIRE study in which new risk prediction models for IFD are developed, validated, and evaluated. The case study illustrates that it is necessary but insufficient for new risk models to show reasonable external validity in random, temporal, and geographical validation samples. Before a particular risk model can be recommended for use in clinical practice, in this case to initiate antifungal prophylaxis, it is important to assess the relative cost-effectiveness of alternative risk prediction strategies. In the FIRE study, the incremental costs of the risk assessment strategies compared with no risk assessment strategy were positive for most of the strategies other than a few strategies that were cost saving. The small cost differences reflect the low unit costs of the prophylaxis, which was the nonproprietary form of fluconazole (400 mg per day). Each of the risk assessment strategies had QALY gains compared with current practice, but these were small and it could be argued that these differences were less than the minimally important difference in QALY gains [53]. The CEA that incorporated these

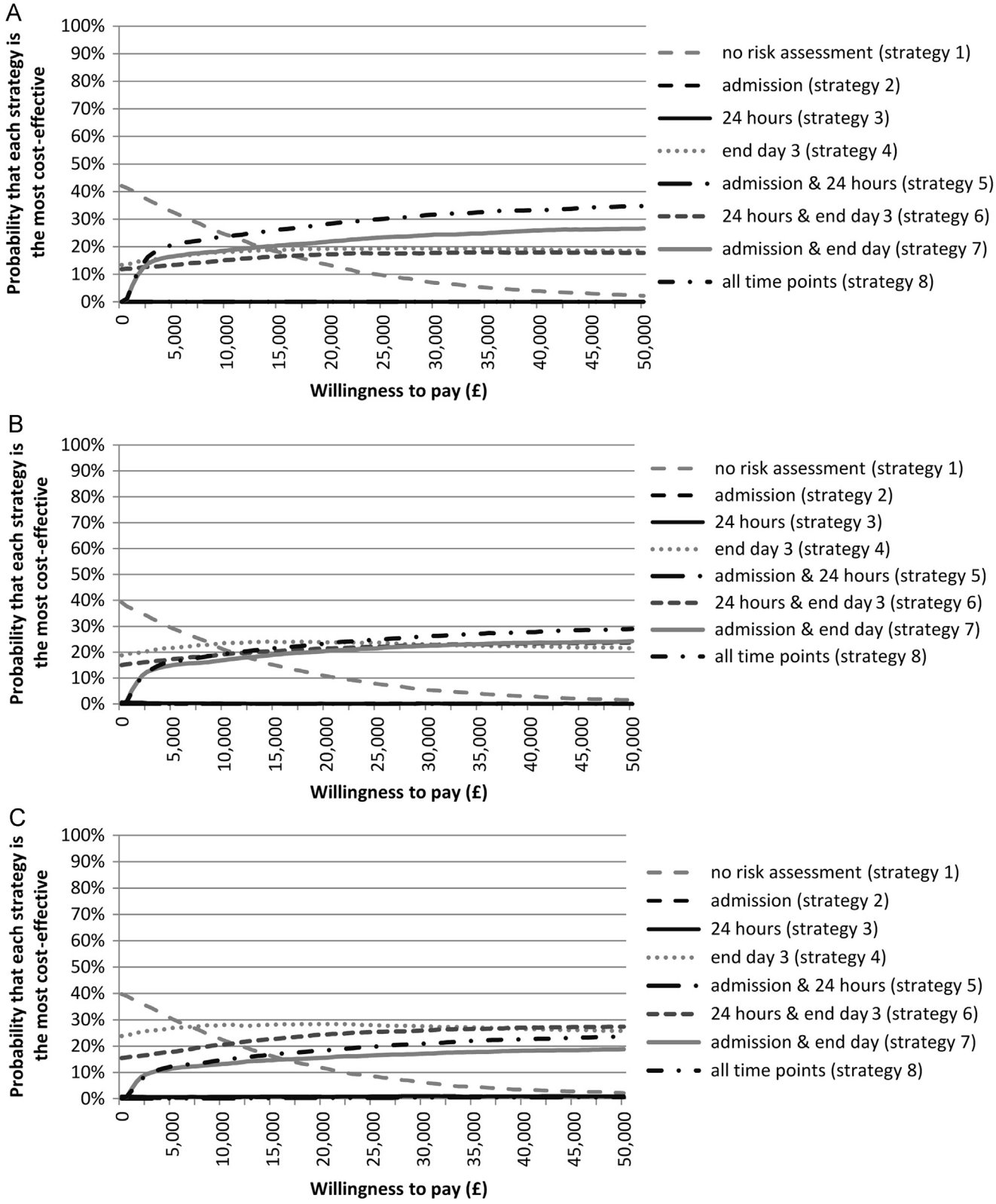


**Fig. 2 – (A) Cost-effectiveness acceptability curves at risk threshold of 0.5%. (B) Cost-effectiveness acceptability curves at risk threshold of 1%. (C) Cost-effectiveness acceptability curves at risk threshold of 2%.**

small differences in cost and QALY across strategies reported that risk assessment and prophylaxis at all time points, and at the end of calendar day 3, was the strategy most likely to be cost-effective when the risk threshold was 0.5% and 1% or 2%, respectively.

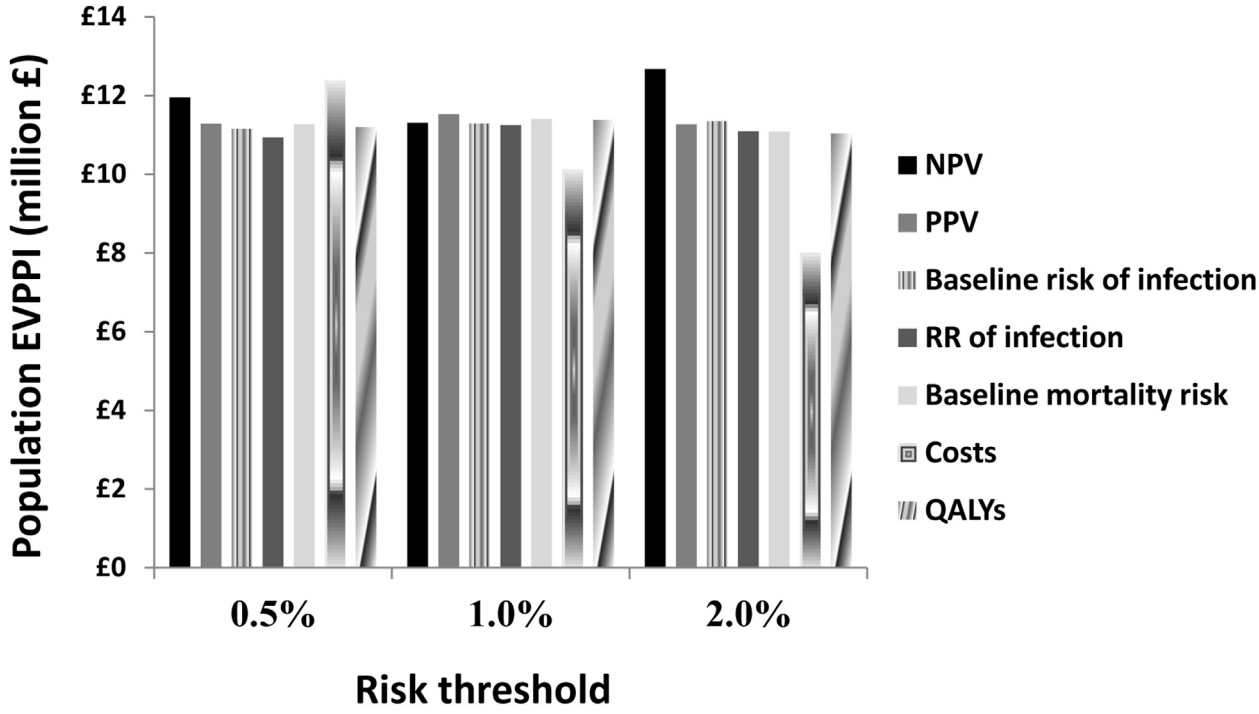


**Fig. 3 – Population EVPPI at £20,000 per QALY, for alternative risk thresholds. EVPPI, expected value of partial perfect information; NPV, negative predictive value; PPV, positive predictive value; QALY, quality-adjusted life-year.**

The integrated approach highlights that the best-fitting risk prediction approach may not be the most cost-effective. For example, a risk prediction strategy at the end of 24 hours provides the best discrimination (according to the c statistic and Brier's score from combined validation), but this strategy is not the most cost-effective. An important aspect of the integrated approach is VOI, as the FIRE study highlights that even with a risk prediction model that performs reasonably well, the decision uncertainty may be such that further research to estimate parameters such as the NPV and the PPV with greater precision can be worthwhile. Hence, an integrated approach to the development and evaluation of such risk models can be useful to inform both clinical decisions and future research priorities.

This integrated approach to the development and evaluation of risk prediction approaches within the same study has several advantages over the more common approach whereby a risk prediction model is developed in one study and used in a subsequent CEA [11–15]. First, the integrated approach starts by defining the clinical decision problem and uses this to stipulate which risk prediction model(s) is required. It can facilitate the development of risk prediction models for specific treatment modalities or patient subgroups, relevant to the decision problem. Second, the article follows guidelines for the development and validation of a risk model [34,36] and extends previous work [16,17] in this area by adopting a more extensive validation of the risk prediction model than that routinely undertaken and by evaluating the accompanying decision uncertainty.

This article extends previous uses of risk prediction models in CEA [9,11–17], and while this integrated approach has been considered in the specific example of the FIRE study, it can be applied more generally. Possible examples include evaluating strategies for preventing emergency admissions in high-risk patients [1], initiating high-cost treatments for patients with hepatitis C infection, initiating herceptin treatment in subgroups of patients with breast cancer [54–56], and targeting treatment for severe sepsis in intensive care [46,57]. While in each of these areas prognostic models exist, it would be useful to assess whether the additional costs of categorizing patients for receiving treatment according to new prognostic markers (such as IL28 for hepatitis C) are worthwhile.

This study has some limitations. First, the approach is illustrated and no consideration is given to the prevention of onward transmission. In the FIRE study, when the risk prediction models included an additional variable to indicate the presence of another patient with IFD in the critical care unit at the same time, this did not improve model performance, which suggests that onward transmission was not a major factor. Second, the CEA results do not apply to the small minority of patients in critical care units currently prescribed antifungal prophylaxis according to clinical judgment (1% of eligible admissions in the FIRE study). Even when patients who received antifungal prophylaxis were included in the study, the baseline risk of IFD remained low (<0.5%). Third, the integrated approach allowed the development and validation of risk prediction models according to each clinical decision time point, but validation was performed only at each single time point whereas the decision problem included risk assessment at multiple time points.

Last, the decision analytical model ignored any impact of the increased use of prophylaxis on antifungal resistance. Including the effects of resistance on the costs and health outcomes of future patients would reduce the relative cost-effectiveness of the risk assessment strategies compared with current practice. In particular, while at the lower risk threshold (0.5%), the most cost-effective strategy was to assess risk, and to provide prophylaxis for those patients whose predicted risk exceeded the threshold, this would require around one-third of patients to receive antifungal prophylaxis. Providing antifungal prophylaxis to a high proportion of critically ill patients raises concerns that future patients may then be resistant to fluconazole. The possible consequences of resistance to antifungal prophylaxis could include increased lengths of stay in critical care units and in hospitals, the additional diagnostic tests, and treatment costs for patients infected with a resistant

form [58]. A key cost may be that resistance may make current antifungal agents unusable—either because they are no longer effective or because of the fear of resistance. This may then make next-generation antifungals necessary, which has implication on both time and cost of developing next-generation antifungals. Studies to date have not fully assessed the cost of resistance to antifungal prophylaxis. In a related context, Smith and Coast [59] highlighted that published studies underestimated the true costs of antibacterial resistance. Further research is required to consider the full costs of antifungal prophylaxis in terms of the additional burden to future patients whose treatment with antifungal agents becomes inappropriate due to increased resistance [60]. Incorporating these effects of resistance in decision analytic modeling is challenging because it requires estimates of additional parameters, such as the resistance rate, the ensuing effect on morbidity and mortality, and a broader model structure, to consider future populations who may be affected by increased resistance.

The integrated approach proposed offers some interesting avenues for further research. The development of risk prediction methods could be extended to other techniques proposed in the methodological literature such as neural networks, support vector machines, and random forests [61–66]. While there is some evidence from simulation studies that these approaches can improve prediction, they are less transparent [67]. More importantly, such approaches should be subjected to the same scrutiny as the integrated approach presented here.

In conclusion, this article illustrates an integrated approach to developing and evaluating risk models within the same study. This approach can be applied more generally to help decision makers appropriately target treatment and prevention strategies.

## Acknowledgments

We thank Adam Steventon, Noémi Kreif, Karla Diaz-Ordaz, Rosa Legood, Mark Pennington, Rosalba Radice, David Epstein, and Manuel Gomes for their helpful comments. We also thank the FIRE Study Steering Group (B. Riley, R. Barnes, J. Edgeworth, R. Grieve, D. Harrison, M. Jit, C. Kibbler, R. McMullan, K. Rowan, N. Soni, and T. Stambach), the FIRE Study Team at ICNARC, and the patients and staff of participating critical care units.

Source of financial support: The FIRE study was funded by the National Institute for Health Research Health Technology Assessment (NIHR HTA) program (project number 07/29/01) and published in full in *Health Technology Assessment* (2013) [23]. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the HTA program, the NIHR, the National Health Service, or the Department of Health.

## Supplemental Materials

Supplemental material accompanying this article can be found in the online version as a hyperlink at http://dx.doi.org/10.1016/j.jval.2013.09.006 or, if a hard copy of article, at www.valueinhealthjournal.com/issues (select volume, issue, and article).

## REFERENCES

[1] Billings J, Dixon J, Mijanovich T, Wennberg D. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. BMJ 2006;333:327.

[2] Chatterton H, Younger T, Fischer A, Khunti K. Risk identification and interventions to prevent type 2 diabetes in adults at high risk: summary of NICE guidance. BMJ 2012;345:e4624.

[3] Hippisley-Cox J, Coupland C, Vinogradova Y, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ 2008;336:1475–82.

[4] National Institute for Health and Care Excellence. Guide to the Methods of Technology Appraisal. London, UK: National Institute for Health and Care Excellence, 2008.

[5] Siontis GC, Tzoulaki I, Siontis KC, Ioannidis JP. Comparisons of established risk prediction models for cardiovascular disease: systematic review. BMJ 2012;344:e3318.

[6] Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: the Framingham Study. Am J Cardiol 1976;38:46–51.

[7] Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat 1992;22:207–19.

[8] Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system: risk prediction of hospital mortality for critically ill hospitalized adults. Chest 1991;100:1619–36.

[9] Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. JAMA 1993;270:2957–63.

[10] Lloyd-Jones DM. Cardiovascular risk prediction: basic concepts, current status, and future directions. Circulation 2010;121:1768–77.

[11] Owens DK, Sanders GD, Heidenreich PA, et al. Effect of risk stratification on cost-effectiveness of the implantable cardioverter defibrillator. Am Heart J 2002;144:440–8.

[12] Sheldon R, O'Brien BJ, Blackhouse G, et al. Effect of clinical risk stratification on cost-effectiveness of the implantable cardioverter-defibrillator: the Canadian implantable defibrillator study. Circulation 2001;104:1622–6.

[13] Grieve R, Hutton J, Green C. Selecting methods for the prediction of future events in cost-effectiveness models: a decision-framework and example from the cardiovascular field. Health Policy 2003;64:311–24.

[14] Williams C, Brunskill S, Altman D, et al. Cost-effectiveness of using prognostic information to select women with breast cancer for adjuvant systemic therapy. Health Technol Assess 2006;10:iii–iv, ix–xi, 1–204.

[15] Longworth L, Young T, Buxton MJ, et al. Midterm cost-effectiveness of the liver transplantation program of England and Wales for three disease groups. Liver Trans 2003;9:1295–307.

[16] Henriksson M, Palmer S, Chen R, et al. Assessing the cost effectiveness of using prognostic biomarkers with decision models: case study in prioritising patients waiting for coronary artery surgery. BMJ 2010;340:b5606.

[17] Rapsomaniki E, White IR, Wood AM, Thompson SG. A framework for quantifying net benefits of alternative prognostic models. Stat Med 2012;31:114–30.

[18] Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Making 2004;24:207–27.

[19] Gudlaugsson O, Gillespie S, Lee K, et al. Attributable mortality of nosocomial candidemia, revisited. Clin Infect Dis 2003;37:1172–7.

[20] Pittet D, Li N, Woolson RF, Wenzel RP. Microbiological factors influencing the outcome of nosocomial bloodstream infections: a 6-year validated, population-based model. Clin Infect Dis 1997;24:1068–78.

[21] Rentz AM, Halpern MT, Bowden R. The impact of candidemia on length of hospital stay, outcome, and overall cost of illness. Clin Infect Dis 1998;27:781–8.

[22] Playford EG, Webster AC, Sorrell TC, Craig JC. Antifungal agents for preventing fungal infections in non-neutropenic critically ill and surgical patients: systematic review and meta-analysis of randomized clinical trials. J Antimicrobial Chem 2006;57:628–38.

[23] Harrison D, Muskett H, Harvey S, et al. Development and validation of a risk model for identification of non-neutropenic, critically ill, adult patients at high risk of invasive Candida infection: the Fungal Infection Risk Evaluation (FIRE) study. Health Technol Assess 2013;17(3):1–156.

[24] Health Protection Agency. Fungal Diseases in the UK – The Current Provision of Support for Diagnosis and Treatment: Assessment and Proposed Network Solution. London: Health Protection Agency, 2006.

[25] Piarroux R, Grenouillet F, Balvay P, et al. Assessment of preemptive treatment to prevent severe candidiasis in critically ill surgical patients. Crit Care Med 2004;32:2443–9.

[26] Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.

[27] Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? BMJ 2009;338:b375.

[28] Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiol 2010;21:128–38.

[29] Altman D, Royston P. What do you mean by validating a prognistic model? Stat Med 2000;19:453–73.

[30] Allen E, Patel K, Rowan KM, et al. Development and validation of risk models for identification of non-neutropenic, critically ill, adult patients at high risk of invasive Candida infection: the Fungal Infection Risk Evaluation (FIRE) study. (Draft paper - personal communication). 2012.
[31] Muskett H, Shahin J, Eyres G, et al. Risk factors for invasive fungal disease in critically ill adult patients: a systematic review. Crit Care 2011;15:R287.
[32] Bewick V, Cheek LJB. Statistics review 14: logistic regression. Crit Care Med 2005;9:112–8.
[33] Austin PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. J Clin Epidemiol 2004;57:1138–46.
[34] Harrell FE Jr, Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. JAMA 1982;247:2543–6.
[35] King G, Zeng L. Logistic regression in rare events data. Pol Analysis 2001;9:137–63.
[36] Steyerberg E. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer, 2009.
[37] Huber PJ. The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions. Berkeley: University of California Press, 1967.
[38] White H. Maximum likelihood estimation of misspecified models. Econometrica 1982;50:1–25.
[39] Brier GW. Verification of forecasts expressed in terms of probability. Monthly Weather Rev 1950;78:1–3.
[40] Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. JASA 1983;78:316–31.
[41] Curtis L. Unit Costs of Health & Social Care 2010. Canterbury, UK: Personal Service Research Unit, University of Kent, 2010.
[42] Joint Formulary Committee. British National Formulary (online). London: BMJ Group and Pharmaceutical Press. Available from: http://www.bnf.org/bnf/index.htm. [Accessed March 25, 2013].
[43] Davies A, Ridley S, Hutton J, et al. Cost effectiveness of drotrecogin alfa (activated) for the treatment of severe sepsis in the United Kingdom. Anaesthesia 2005;60:155–62.
[44] Department of Health. NHS Reference Costs 2009-2010. London, UK: Department of Health, 2011.
[45] Cuthbertson BH, Roughton S, Jenkinson D, et al. Quality of life in the five years after intensive care: a cohort study. Crit Care 2010;14:R6.
[46] Green C, Dinnes J, Takeda A, et al. Clinical effectiveness and cost-effectiveness of drotrecogin alfa (activated) (Xigris) for the treatment of severe sepsis in adults: a systematic review and economic evaluation. Health Technol Assess 2005;9:1–126, iii–iv.
[47] Wright JC, Plenderleith L, Ridley SA. Long-term survival following intensive care: subgroup analysis and comparison with the general population. Anaesthesia 2003;58:637–42.
[48] Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey. Centre for Health Economics Discussion Paper 1995;138:1–19.
[49] Briggs A, Sculpher M, Claxton K. Decision Modelling for Health Economic Evaluation. Oxford: Oxford University Press, 2006.
[50] Lambert PC, Sutton AJ, Burton PR, et al. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. Stat Med 2005;24:2401–28.
[51] Soares MO, Welton NJ, Harrison DA, et al. An evaluation of the feasibility, cost and value of information of a multicentre randomised controlled trial of intravenous immunoglobulin for sepsis (severe sepsis and septic shock): incorporating a systematic review, meta-analysis and value of information analysis. Health Technol Assess 2012;16:1–186.
[52] Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ 1999;18:341–64.
[53] Walters SJ, Brazier JE. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. Health Qual Life Outcomes 2003;1:4.
[54] National Institute for Health and Care Excellence. Trastuzumab for the Adjuvant Treatment of Early Stage HER2-Positive Breast Cancer. London, UK: National Institute for Health and Care Excellence, 2007.
[55] National Institute for Health and Care Excellence. Boceprevir for the Treatment of Genotype 1 Chronic Hepatitis C. London, UK: National Institute for Health and Care Excellence, 2012.
[56] National Institute for Health and Care Excellence. Telaprevir for the Treatment of Genotype 1 Chronic Hepatitis C. London, UK: National Institute for Health and Care Excellence, 2012.
[57] Sadique MZ, Grieve R, Harrison DA, et al. Is Drotrecogin alfa (activated) for adults with severe sepsis, cost-effective in routine clinical practice? Crit Care 2011;15:R228.
[58] McGowan JE Jr. Economic impact of antimicrobial resistance. Emerg Infect Dis 2001;7:286–92.
[59] Smith R, Coast J. The true cost of antimicrobial resistance. BMJ 2013;346:f1493.
[60] Cataldo MA, Petrosillo N. Economic considerations of antifungal prophylaxis in patients undergoing surgical procedures. Ther Clin Risk Manag 2011;7:13–20.
[61] Breiman L. Bagging predictors. Mach Learn 1996;24:123–40.
[62] Dybowski R, Gant V. Clinical Applications of Artificial Neural Networks. Cambridge: Cambridge University Press, 2001.
[63] Ripley B. Pattern Recognition and Neural Networks. Cambridge: Cambridge University Press, 1996.
[64] Russel S, Norvig P. Artificial Intelligence: A Modern Approach. Boston: Pearson, 2010.
[65] Van Looy S, Verplancke T, Benoit D, et al. A novel approach for prediction of tacrolimus blood concentration in liver transplantation patients in the intensive care unit through support vector regression. Crit Care 2007;11:R83.
[66] Verplancke T, Van Looy S, Benoit D, et al. Support vector machine versus logistic regression modeling for prediction of hospital mortality in critically ill patients with haematological malignancies. BMC Med Inform Decis Mak 2008;8:56.
[67] Sampson DL, Parker TJ, Upton Z, Hurst CP. A comparison of methods for classifying clinical samples based on proteomics data: a case study for statistical and machine learning approaches. PloS One 2011;6:e24973.