The evidential basis of decision making in plant disease management

Hughes, G

Published in:
Annual Review of Phytopathology

DOI:
10.1146/annurev-phyto-080516-035342

First published: 10/05/2017

Document Version
Peer reviewed version

The Evidential Basis of Decision Making in Plant Disease Management

Gareth Hughes

Crop and Soil Systems Research Group, SRUC, Edinburgh EH9 3JG, UK; email: gareth.hughes@sruc.ac.uk

TABLE OF CONTENTS

Keywords

risk, risk factors, Bayes' rule, likelihood ratios, weight of evidence

Abstract

The evidential basis for disease management decision making is provided by data relating to risk factors. The decision process involves an assessment of the evidence leading to taking (or refraining from) action on the basis of a prediction. The primary objective of the decision process is to identify – at the time the decision is made – the control action that provides the best predicted end-of-season outcome, calculated in terms of revenue or another appropriate metric. Data relating to disease risk factors take may a variety of forms (e.g., continuous, discrete, categorical) on measurement scales in a variety of units. Log_{10}-likelihood ratios provide a

principled basis for the accumulation of evidence based on such data, and allow predictions to be made via Bayesian updating of prior probabilities.

INTRODUCTION

The 1959 publication of *The Integrated Control Concept* (77) established thresholds as part of the apparatus of decision making in crop protection at the field scale. The 'economic injury level' (EIL) was defined as the lowest population density that will cause an amount of crop injury that justifies the cost of artificial control measures; the 'economic threshold' (ET) as the population density at which control measures should be used to prevent an increasing pest population from reaching the EIL (77). Because these concepts were originally described in relation to the management of arthropod pests, pest population density was the natural metric in which to denominate threshold levels. As conceived, if pest population density exceeds the ET, a decision to initiate artificial control measures (in order to prevent to pest population reaching the EIL) is justifiable on the basis of a calculation comparing the predicted revenues from treated and untreated crops. The analysis implicitly assumes application of control measures at a fixed, pre-specified rate.

To put these ideas into practice, we first need to describe the relationship between yield reduction and the intensity of harmful organisms (15). Such a crop loss model characterizes the effect of the harmful organism population of concern on untreated host crops, providing a basis for calculating an EIL and a corresponding ET. Then, operationally, an appropriate sampling methodology is required in order to classify the level of the harmful organism population of concern relative to the ET. In the context of decision making in crop protection, acceptance sampling methods borrowed from statistical quality control (64) have been adopted by economic entomologists for the purpose of monitoring pest populations and classifying population densities relative to thresholds (8). If the threshold levels of harmful organisms are calculated in

terms of pest population density, the decision process is facilitated if sampling provides evidence denominated in the same natural metric.

Looking at this background more specifically in relation to decision making for plant disease management (and restricting attention here mainly to the characteristics of pathosystems concerning annual field crops), a few points of detail need attention. In phytopathology, population density is not often an appropriate metric for assessing disease, denominating thresholds and calibrating crop loss models. Instead, disease intensity is usually assessed in terms of severity (the area of plant tissue affected by disease expressed as a proportion or percentage of the total area) or incidence (the number of plant units infected, expressed as a proportion or percentage of the total number of units assessed). Relationships between incidence and severity (62, 74) are sometimes useful for converting between measures of intensity. Many phytopathological crop loss models are denominated in terms of disease severity or disease incidence (17, 50), and the corresponding ETs and EILs are then expressed in the same measure of intensity. Thus severity and incidence have often been used as natural metrics for evidence in the decision process for plant disease management, although it is clear that this represents a simplification of the interaction between host and pathogen (58).

According to this approach, sampling is required to obtain the evidence base of disease intensity needed for the decision process. For crop protection decision making, classification of severity or incidence is usually more appropriate than estimation (65), so acceptance sampling is again the appropriate method for monitoring disease intensity in relation to thresholds (44, 60). Operating characteristic (OC) curves are instrumental in establishing and evaluating the performance of such sampling schemes. In acceptance sampling, an OC curve characterizes the probability that a sample has been drawn from a population that is regarded as acceptable (i.e., in

the present context, not exceeding the adopted threshold level at which control measures are deemed necessary) as a function of the true population value. The OC curve decreases as the true population value increases. The resulting decision is binary; control measures are either initiated, or not initiated. Since the classification on which this decision is based depends on sampling, we must recognise that a crop may be misclassified. Acceptance sampling methodology is less familiar in phytopathology than in economic entomology, but there are some useful examples (31, 32, 48, 49, 80).

As originally described, the integrated control concept based on the EIL and the ET is not without shortcomings (8, 72, 89). Like Binns et al. (8) we recognise the usefulness of thresholds as guidelines in crop protection decision making, while also acknowledging that economic advantage may not be the only management objective driving those decisions. Although crop protection decisions may not necessarily be made by explicit reference to pre-determined economic thresholds, it nevertheless seems a reasonable proposition that an individual's decision to take or not to take an action in some way represents reference to a – perhaps individual, perhaps implicit – threshold. All actionable decisions are liminal.

Here, our main focus will be on the evidential basis for decision making, rather than the thresholds themselves. Particularly (although not exclusively) of interest in this context is evidence related to plant disease management that is multi-faceted, such that disease intensity may only be one of a number of factors contributing to the decision process.

RISK

We all may invoke the idea of risk from time to time in everyday life, with various shades of meaning usually related informally to the chance of some undesirable circumstance involving injury, harm or loss. Even when it comes to formal descriptions of risk, the literature provides us with a variety of definitions. For a technical discussion of risk, therefore, some care is required in communicating just what is meant.

For the U.K., the National Risk Register (11, 12) uses a simple diagram to compare a range of events that might have a major impact on all (or at least significant parts) of the country. A two-dimensional array is depicted with increasing 'relative likelihood' or 'relative plausibility of occurring' on the horizontal axis and increasing 'relative impact' on the vertical axis (the axes may have continuous or ordinal categorical scales). The events of concern may then be placed at an appropriate position in the array, on the basis of the available evidence. The horizontal axis is, in essence, a scale representing the probability of an event for which the outcome is undesirable. In epidemiology, the term 'risk' is used to refer to such a probability (61, 75). For terminological clarity, we shall where necessary in this review refer to 'epidemiological risk' when it is strictly this probability that is under discussion. It is convenient that the literature on statistical quality control (64) also uses a probabilistic definition of risk, so there is no disjunct between the discussion of epidemiological risk based on evidence from acceptance sampling for disease intensity and on evidence from a wider range of factors. Even when not expressed strictly in probabilistic terms as epidemiological risk, informal expressions of risk in the context of crop protection are usually associated with circumstances indicative of an increased chance of the need for a control intervention. Either way, we do not measure risk directly, but via identification of those circumstances and quantification of the effect of their occurrence.

This brings us to risk factors – a term originally used in clinical epidemiology for behaviors and biomarkers identified as predictors of disease (55). Note that this epidemiological terminology does not imply that observed data for such predictors have any particular statistical characteristics. Thus in phytopathology, risk factors may be continuous (e.g., many meteorological variables), discrete (including binary) (e.g., years since previous host crop, disease outbreak or not during the preceding year), ordinal categorical (e.g., varietal disease resistance rating) or nominal categorical (e.g., variety). In addition, the scales on which risk factors are measured may sometimes be discretized or categorized to facilitate description or analysis. The identification of disease risk factors and quantification of their evidential value has become a fundamental aspect of modern epidemiology. However, a formal description of risk is not necessarily a prerequisite for the analysis of risk factors. For example, we may think of – and refer to – the intensity of harmful organisms (as outlined above in relation to crop loss) as a risk factor, without the need for a strictly probabilistic interpretation of risk. Particularly if a decision process is specified in terms of a single risk factor, it may be appropriate to analyse the evidential basis for prediction of crop loss using the natural measurement scale of that factor.

As mentioned above, the analysis underlying the economic threshold concept is based on the assumption that it is a discrete choice threshold: following a decision to use control measures, application is made at a predetermined dose rate. Greater flexibility is possible. For example, in the UK, the HGCA (Home Grown Cereals Authority, now under the auspices of the AHDB (Agriculture and Horticulture Development Board)) has for a number of years conducted research aimed at characterizing appropriate fungicide doses, facilitating the identification of a dose rate that maximises economic benefit given the degree of disease risk, by balancing the effectiveness of disease control against the cost of fungicide (66, 67).

Whatever the operational details of implementing a crop protection decision process, we can see that it involves taking (or refraining from) action on the basis of a prediction. The evidential basis for the decision is provided by risk factors assessed during a period leading up to the time that the decision is made. The objective of the decision process is to identify – at the time the decision is made – the control action that provides the best predicted end-of-season outcome, calculated in terms of revenue or whatever is deemed the appropriate metric.

IDENTIFICATION AND ANALYSIS OF RISK FACTORS

In many cases, decision makers are familiar with the likely risk factors for common diseases. In the U.K., for example, the HGCA guides for cereals and oilseeds include lists of identified risk factors for the most common diseases (4, 5, 6) based on data either from new purpose-designed experiments or from retrospective examination of previous experiments. Further to the identification of risk factors, the next steps are formal analysis of their capacity to predict crop loss, and incorporation in a predictive system for disease management. In essence, identification and analysis of risk factors is the activity that underlies the development of crop loss models.

Identification of Single Factors

Many single-factor crop loss models constitute a relationship between yield reduction and disease intensity (17, 50). However, risk factors assessed as a basis for predicting crop loss need not necessarily be direct assessments of disease intensity. For example, Paveley et al. (69) studied determinants of fungicide spray decisions for wheat (referred to as ‘risk determinants’).

In this case, measures of disease-induced loss of green leaf area index were found to be more consistent predictors of yield loss across sites and seasons than measures of disease intensity.

Identification of Multiple Factors

Sometimes more than one risk factor is associated with crop loss. For example, Mila et al. (63) studied risk factors for panicle and shoot blight of pistachio (caused by *Fusicoccum* sp.), an economically important disease in California. The aim of the experimental study was to quantify the effects of latent infection, temperature, precipitation, and irrigation systems on disease severity, with the ultimate objective of providing a basis for development of a risk assessment system for panicle and shoot blight of pistachio to guide growers in their efforts to control the disease. Note in passing that this kind of study is not necessarily dependent solely on field experimentation. Survey data may be useful, particularly in the case of infrequent disease threats (57). In veterinary epidemiology, elicitation of expert opinion has successfully been used in the identification of risk factors (37, 85).

Identification of Multiple Factors Plus Risk Assessment

Once multiple risk factors have been identified, developing a system for risk assessment requires quantitative analysis. As part of this analysis some consideration must be given to the appropriate metric for evidence, when different risk factors have different measurement scales. Studies of spotted wilt of peanut (caused by *Tomato spotted wilt virus*) (18) leading to the development of a method of risk assessment contributing to integrated management of the

disease provide an example (9). The risk assessment system comprises a table of risk factors (in this case, peanut production practices) with 'risk index points' given for each level of each factor. The relative weighting of risk points was established by reference to available research and, in some cases, the authors' expert opinions. For operational risk assessment, the appropriate points are accumulated for each factor for the crop in question, and the resulting points total is a predictor of risk of losses to spotted wilt disease, categorized as low/medium/high. Graphical plots of % spotted wilt severity against accumulated risk index points are described by linear relationships. These relationships are, in effect, single factor crop loss models, where the risk index has subsumed the set of individual risk factors. However, there is a good reason to maintain the tabular format of risk assessment system as the operational version. The table supports decision making at the tactical level of protecting a particular crop, and also contributes to decision making at the strategic level, by raising awareness of the most important factors by means of which risk may be modified. Surveys provide evidence of how the risk index has mediated changes in peanut production practices in Georgia, relating to cultivar selection, planting date, row pattern, tillage, and insecticide selection (9).

Developing a system for risk assessment does not necessarily require that levels of the factors are characterized. For example, studies of potato late blight (PLB, caused by *Phytophthora infestans*) in south-central Washington led to the development of a regional risk assessment based on a data set including continuous (meteorological) factors (53).

Risk Assessment in a New Location

Sometimes, a risk assessment system developed for a disease problem in one area may be used in a new area. For example, Duttweiler et al. (20) studied the performance of a risk assessment system for a disease complex of apple (sooty blotch and flyspeck) developed in the south-eastern United States, when used in the upper-midwest United States. Sporadic disease control failures in the new location led to the need to determine whether modification of the system could provide more reliable risk assessment. Essentially, using a pre-existing system in a new location puts us back in the situation where we are familiar with the likely risk factors but require further formal analysis of data obtained in the new location to corroborate (or otherwise) the extent to which their capacity to predict crop loss is transferrable. We are back to the identification and analysis of risk factors.

Logistic Regression

As might easily be appreciated, various statistical methods have been deployed in studies of identification and weighting of risk factors as the basis for development of predictive systems in disease management. It is beyond the scope of the present review to describe all these methods, and interested readers are referred to the original sources for details. It is particularly notable, however, that logistic regression has had frequent application in the analysis of disease risk factors. Consider the following scenario. For a particular pathosystem, a list of putative risk factors is prepared. For a set of crops, all untreated for the disease in question, the listed risk factors are observed and recorded in each crop. At the end of the season, disease intensity (or an equivalent metric in which the EIL is specified) is recorded in each crop. Crops for which the

EIL has been exceeded are noted (as shorthand, crops in this group are referred to here as 'cases') and crops for which the EIL has not been exceeded similarly noted (and referred to here as 'controls'). The designation as case (recorded as 1 for data analysis) or control (recorded as 0) is regarded as the gold standard. The practical problem is that these gold standard determinations cannot be used for decision making in practice, because by the time they are made, it is too late for preventative crop protection measures to be taken. Instead, we require a prediction of need for treatment that can be made earlier, at the ET. The statistical problem is then to model a binary outcome variable (case or control) as a function of one or more explanatory variables. In logistic regression, the logit transformation of a probability *p* (epidemiological risk; which depending on context may refer to the probability of disease outbreak or of need for a control intervention, for example) is modelled as a linear function of explanatory variables (16), here denoted *f*(risk factors):

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = f(\text{risk factors}),$$

and noting that the inverse function of logit(●) is the logistic function, we have:

$$p = \frac{1}{1+e^{-f(\text{risk factors})}}.$$

Relationships of this form are essentially crop loss models in which the downside is expressed in terms of epidemiological risk rather than yield reduction.

Where more than one risk factor is associated with crop loss, the parameter estimation process initially involves identification of those factors that, on the basis of the data, merit inclusion in the model. Note that the composition of *f*(risk factors) is not restricted to main effects; interaction

terms may also be included in the model. In addition to identification of factors associated with epidemiological risk, we wish to avoid over-counting of evidence resulting from inclusion of two or more factors that are associated with crop loss but are non-independent. To this end, most statistical software for multiple logistic regression analysis implements forwards-inclusion/backwards-elimination techniques that allow stepwise model-fitting and model-checking. This process may result in more than one candidate model for further consideration. Candidate models will then usually undergo some form of goodness-of-fit assessment, possibly involving cross-validation on data not used in model-fitting. This is beyond the scope of the present review; our goal here is to characterize the way in which we can quantify evidence related to risk factors so as to be able – in future – to make predictions of the status of crops on the basis of observations of those risk factors. Most phytopathological applications of logistic regression analysis in this context are for binary logistic regression, where two classes are available for prediction of the status of a crop (i.e., classification is made relative to a single threshold). Ordinal logistic regression can be used if more than two classes of prediction are needed (68). As a rule of thumb, a decision process requires no more classes of prediction than the number of different actions available to the decision maker.

Binary Logistic Regression with a Single Risk Factor

Fabre et al. (23) and Pethybridge et al. (71) have provided examples of epidemiological risk modelled by logistic regression on a single risk factor. The resulting analyses were depicted by sigmoid curves on graphs with epidemiological risk on the vertical axis and evidence related to the single factor (in each case, an appropriate measure of intensity) on the horizontal axis.

Hughes et al. (47) presented similar examples, and placed the analysis in the wider context of crop loss assessment models.

Binary Multiple Logistic Regression with Risk Points

Here, the paradigmatic example is provided by the pioneering work of Jonathan Yuen, Eva Twengström and associates at the Swedish University of Agricultural Sciences (82, 88). Risk factors for Sclerotinia stem rot (caused by *Sclerotinia sclerotiorum*) on oilseed rape in east-central Sweden were identified. Epidemiological risk was modelled by logistic regression on a set of risk factors comprising only main effects. A simple risk point scale for the levels of each risk factor was devised by making minor arithmetic adjustments to the estimated regression coefficients. A graph of end-of-season disease incidence against accumulated risk points, based on data from crops not used for model fitting, illustrated a test of the model. Risk points constitute a metric for evidence that can be summed to provide a points total for a crop. In operation, predictions of the need for a control intervention are made by comparing the risk points accumulated for a crop with a threshold points score. Risk points provide a common currency for the combination of evidence from risk factors measured in different units. A published table of risk points (9, 82) also illustrates for decision makers the relative importance of the different risk factors in a decision process.

Binary Multiple Logistic Regression with Continuous Variables

The example here relates to meteorological risk factors associated with epidemics of Fusarium head blight of wheat (FHB, caused by *Fusarium graminearium*) in three different wheat-production regions of the U.S.A. (19). The FHB pathosystem is complicated, involving a number of pathogens of varying geographical relevance. FHB is important not only because of the resulting yield reduction, but also because it may cause contamination of grain by mycotoxins. A review of predictive models for FHB covers aspects of both risk and geographical variation (73). In the U.K., a points-based risk model is used to identify the appropriate agronomy for minimising risk of mycotoxins (3, 21, 22). Here, the example concerns development of risk models for disease management (19). The starting point was a list of weather variables (including interactions) potentially of value in predicting epidemics of FHB. Logistic regression equations for a set candidate models were estimated and model performance assessed. Such models have application in epidemic prediction; important for farmers in relation to the need for timely disease management intervention. For a multiple logistic regression model based on continuous data, the common-currency metric for evidence from a combination of risk factors – the indicator of the risk of an epidemic – is logit(p) on the left-hand-side of the model equation. When models include continuous variables and their interactions, a published table of risk points is not a practical proposition.

Thresholds

Logistic regression analysis models epidemiological risk as a function of risk factors, but – unlike discriminant function analysis – leaves open the choice of threshold level of evidence

related to risk factors that is used in a decision process. The threshold selected for use with a particular model can be adjusted, based on the risk attitudes of the user and their tolerance for different types of errors (1). This flexibility is one of the reasons that Johnson et al. (52) elected to use logistic regression in their further studies of potato late blight forecasting models. Choosing a threshold is not dealt with in detail here. Swets et al. (79) provide an introduction to receiver operating characteristic (ROC) analysis, widely adopted in phytopathology as the basis for selection of an appropriate operational threshold level in a binary decision process. Incidentally, note also that in an earlier article Swets (78) discussed the relationship between the OC curve and the ROC curve. Yuen et al. (88) and Twengström et al. (82) show ROC curves for their logistic regression models for Sclerotinia stem rot, and Madden (59) shows the ROC curve for FHB logistic regression 'Model B' of De Wolf et al. (19).

BAYES' RULE

Recall the scenario outlined above, whereby for a particular pathosystem, data are collected from a number of untreated crops for logistic regression analysis of a set of risk factors. Crops in the data set are classified by their actual status as cases (where the EIL has been exceeded) or controls (where the EIL has not been exceeded). Following logistic regression analysis, evidence related to risk factors may be quantitatively summarized for each crop, for example in terms of risk points, or in terms of logit(p). Then, after an appropriate operational threshold has been selected, crops in the data set can also be classified by their predicted status as exceeding the threshold or not exceeding the threshold. The data set has thus been cross-classified. A summary of such a cross-classification is provided by a 2×2 table (see Table 2 in (40) for example). We

refer to this table as a prediction-realization table. In clinical epidemiology, this table characterizes a binary diagnostic test. In phytopathology, the prediction-realization table characterizes a binary predictor on the basis of a quantitative summary of the evidence, together with an operational threshold that defines the quantity of evidence required for predictions of an epidemic or of need for a control intervention. Typically, the cross-classification indicates the imperfect nature of the predictions. Some crops that are controls provide enough evidence for an epidemic prediction or a prediction of need for a control intervention. Some crops that are cases provide insufficient evidence for an epidemic prediction or of need for a control intervention.

A prediction-realization table may be used to show either the cross-classified frequencies of predictions and realizations, or the estimated probabilities obtained by normalization of those frequencies. Up to this point we have managed with a minimum of notation, but further analysis of the evidential basis of decision making now requires some symbolic representation of the quantities involved. To this end, we refer to a previous review covering decision making in epidemiology by Madden (59) which included a table listing some of the notation and terminology used in phytopathological applications of decision theory. In the present review we shall as far as possible use a notation close to Madden's, which ought to assist readers who wish to refer to both articles. Actual status of crops in one of two groups (denoted here E_j, j=1,2) is determined by reference to the EIL: those where the EIL was exceeded (cases, denoted E_1) and where the EIL was not exceeded (controls, denoted E_2). Predicted status of crops in one of two groups (denoted here P_i, i=1,2) is determined by reference to the operational threshold: P_1 denotes that the predicted status of a crop is of an epidemic, or need for treatment; P_2 denotes that the predicted status is of a non-epidemic or no need for treatment. Table 1 shows a

normalized 2×2 prediction-realization table in this generic notation. From Table 1 we can obtain the conditional probabilities $\Pr(P_i|E_j)$ and $\Pr(E_j|P_i)$ via Bayes' rule, using:

$$\Pr\left(P_i \cap E_j\right) = \Pr\left(P_i \middle| E_j\right) \cdot \Pr\left(E_j\right) = \Pr\left(E_j \middle| P_i\right) \cdot \Pr\left(P_i\right).$$

Data for a prediction-realization table typically reflect logistic regression analysis of a set of risk factors from a number of untreated crops. The crops are first classified definitively as either case or control, $\Pr(E_j)$. Following logistic regression, frequency distributions of the evidential summary variable may be calculated, separately for cases and controls (see Figure 2 in (82) for example). After an appropriate operational threshold has been selected, crops comprising each of the two separate frequency distributions after classification on the basis of actual status may be further classified by their predicted status as exceeding the threshold or not exceeding the threshold (again see Figure 2 in (82) for example). Thus subsequent to classification based on predicted status the conditional probabilities $\Pr(P_i|E_j)$ may be estimated.

Sensitivity and Specificity

Cases with values of the evidential summary variable exceeding the operational threshold are 'true positives' and controls with values of the indicator variable not exceeding the operational threshold are 'true negatives'. But typically, some cases (as determined by the gold standard) will have values of the evidential summary variable below the threshold, and some controls (as determined by the gold standard) will have values of the evidential summary variable above the threshold. The true positive proportion (*TPP*, sensitivity) expresses the number of true positives as a proportion of the total number of cases. The true negative proportion (*TNP*, specificity)

expresses the number of true negatives as a proportion of the total number of controls. The false negative proportion is $FNP = 1 - TPP$, and the false positive proportion is $FPP = 1 - TNP$. TPP is an estimate of the conditional probability $\Pr(P_1|E_1)$, the probability of an evidential summary value above the threshold, given that the actual status of the crop was a case. TNP is an estimate of the conditional probability $\Pr(P_2|E_2)$, the probability of an evidential summary value at or below the threshold, given that the actual status of the crop was a control. FNP is an estimate of the conditional probability $\Pr(P_2|E_1)$ and FPP is an estimate of the conditional probability $\Pr(P_1|E_2)$. Estimates of the conditional probabilities $\Pr(P_i|E_j)$ are made from data collected while a predictor is being developed, prior to operational implementation.

Sensitivity and specificity represent two kinds of accuracy, respectively, for cases and controls. Sensitivity and specificity are independent of the proportions of cases and controls in a data set and can therefore be viewed as properties of a predictor. In plant disease management, logistic regression analyses leading to an operational predictor often quote the corresponding sensitivity and specificity (e.g., 1, 19, 40, 52, 82, 88).

Likelihood Ratios

In the present context, we are interested in sensitivity and specificity in relation to the evidence provided by a predictor. This is best considered by reference to likelihood ratios: the likelihood ratio of a positive (P_1) prediction $LR(+)$ = sensitivity/(1−specificity) = $\Pr(P_1|E_1)/\Pr(P_1|E_2)$, and the likelihood ratio of a negative (P_2) prediction $LR(-)$ = (1−sensitivity)/specificity = $\Pr(P_2|E_1)/\Pr(P_2|E_2)$. Thus $LR(+)$ tells us the extent to which a P_1 prediction is more likely from E_1 crops as compared with E_2 crops; $LR(-)$ tells us the extent to which a P_2 prediction is less likely

from E_1 crops as compared with E_2 crops. A useful predictor has $LR(+) > 1$ and $0 < LR(-) < 1$; ideally with $LR(+)$ large and $LR(-)$ small. Biggerstaff (7) proposed a simple graphical application of likelihood ratios to facilitate the comparison of predictors; phytopathological applications include (14, 30, 33, 70, 81).

Note that the scenario for data collection outlined above is not the only way to obtain likelihood ratios. Methods for elicitation of subjective likelihood ratios (36) have found application in veterinary epidemiology (38). This methodology typically provides likelihood ratios, but not their component numerators and denominators.

The conditional probabilities $\Pr(E_i|P_j)$ are probability forecasts relating to epidemics or need for treatment. They represent the Bayesian posterior probabilities of an event E_i given data P_j based on:

$$\Pr(E_j|P_i) = \frac{\Pr(P_i|E_j)\cdot\Pr(E_j)}{\Pr(P_i)}.$$

For example:

$$\Pr(E_1|P_1) = \frac{\Pr(P_1|E_1)\cdot\Pr(E_1)}{\Pr(P_1|E_1)\cdot\Pr(E_1)+\Pr(P_1|E_2)\cdot\Pr(E_2)} \qquad (1)$$

is the probability of need for treatment after a prediction of need for treatment and:

$$\Pr(E_2|P_2) = \frac{\Pr(P_2|E_2)\cdot\Pr(E_2)}{\Pr(P_2|E_2)\cdot\Pr(E_2)+\Pr(P_2|E_1)\cdot\Pr(E_1)} \qquad (2)$$

is the probability of no need for treatment after a prediction of no need for treatment (86). To see how these formulae relate to evidence from risk factors, we write *odds*(●) = Pr(●)/(1−Pr(●)); then after some rearrangement we obtain:

$$odds\left(E_1|P_1\right) = odds\left(E_1\right) \cdot LR(+) \qquad (3)$$

from equation (1) and:

$$odds\left(E_2|P_2\right) = odds\left(E_2\right) / LR(-) \qquad (4)$$

from equation (2) (86, 87). In these expressions of Bayes' rule, the posterior conditional odds of an event E_i given the data P_j depends on the prior odds and a likelihood ratio. The likelihood ratio is independent of the prior odds. The prior odds expresses an initial assessment, before use of the predictor; the posterior conditional odds expresses a revision of the initial assessment, after use of the predictor, taking into account evidence related to risk factors. To convert back from odds to probability, Pr(●) = *odds*(●)/(1+*odds*(●)).

In the following section, we will work with a logarithmic form of the odds version of Bayes' rule. As a precursor to that, note that from equation (3) that we can write:

$$\ln\left(odds\left(E_1|P_1\right)\right) = \ln\left(odds\left(E_1\right)\right) + \ln\left(LR(+)\right)$$

from which, following some rearrangement, we obtain:

$$\Pr\left(E_1|P_1\right) = \frac{1}{1+e^{-\left(\ln\left(odds\left(E_1\right)\right)+\ln\left(LR(+)\right)\right)}}$$

which is a logistic equation for the posterior probability $\Pr(E_1|P_1)$. This aside is just meant as a further indication that the methods we have outlined for analysis of the evidential basis of decision making are not unrelated.

WEIGHT OF EVIDENCE

Readers who have seen the 2014 film *The Imitation Game* may not immediately recognise the portrayal of the work of Alan Turing and his group of cryptanalysts at Bletchley Park during WWII as bearing much relation to the process of disease management decision making outlined above. But Turing's group crucially employed Bayesian methods to infer the daily settings of the German naval Enigma machine (35). Indeed, IJ Good – portrayed as a minor character in the film – actually had a major role after WWII in the renaissance of statistical applications of Bayesian methods in decision theory. Turing's group worked with $\log_{10}$-likelihood ratios and it is convenient, for consistency with literature cited, that we do the same from this point on. In this form, we have:

$$\log_{10}\left(odds\left(E_1|P_1\right)\right)=\log_{10}\left(odds\left(E_1\right)\right)+\log_{10}\left(LR(+)\right) \qquad (5)$$

$$\log_{10}\left(odds\left(E_2|P_2\right)\right)=\log_{10}\left(odds\left(E_2\right)\right)-\log_{10}\left(LR(-)\right) \qquad (6)$$

$$\log_{10}\left(odds\left(E_1|P_2\right)\right)=\log_{10}\left(odds\left(E_1\right)\right)+\log_{10}\left(LR(-)\right) \qquad (7)$$

$$\log_{10}\left(odds\left(E_2|P_1\right)\right)=\log_{10}\left(odds\left(E_2\right)\right)-\log_{10}\left(LR(+)\right) \qquad (8)$$

as versions of Bayes' rule. Note that a useful predictor has $\log_{10}(LR(+)) > 0$ (larger values are better) and $\log_{10}(LR(-)) < 0$ (larger negative values are better). Thus an epidemic (or need for treatment) prediction increases the posterior log-odds of an epidemic (or need for treatment) (equation 5) and decreases the posterior log-odds of a non-epidemic (or no need for treatment) (equation 8). A non-epidemic (or no need for treatment) prediction increases the posterior log-odds of a non-epidemic (or no need for treatment) (equation 6) and decreases the posterior log-odds of an epidemic (or need for treatment) (equation 7).

The log-likelihood ratio is referred to as the 'weight of evidence'. Turing's group used the 'ban' as the unit for weight of evidence with base-10 logarithms (35) but this name has fallen out of use, and the unit is now usually referred to as the 'hartley' (Hart) (41). Note in passing that a version of Biggerstaff's likelihood ratios graph (7) for $\log_{10}$-likelihood ratio axes has been described (54, see also 27).

Weight of Evidence in Theory

The use of likelihood ratios as a basis for quantifying evidence related to risk factors in a decision process has not been unreservedly accepted. Zweig & Campbell (90, page 570) wrote: "The conceptual meaning of likelihood ratio is tricky and can be confusing." In a review of the properties of diagnostic tests as used in clinical decision making, Langlotz (56, page 9) wrote: "The likelihood ratio has several properties that limit its usefulness in describing diagnostic examinations" and, among these limitations, "… it is sometimes counterintuitive that the same likelihood ratio causes different absolute changes in probability, depending on the preexamination probability." Zweig & Campbell (90) and Langlotz (56), it should be said, are

not alone in having given voice to such concerns, and as a result, methods have been developed to circumvent perceived difficulties in the application of likelihood ratios to clinical diagnosis while enabling practitioners to take advantage of the benefits (13, 24, 42, 43).

Here we will briefly review the "sometimes counterintuitive" properties of log-likelihood ratios from an information theoretic perspective. For this purpose, we refer to a previously-mentioned study of FHB of winter wheat (19, 59). Readers are encouraged to refer to the original sources for a detailed account of the analytical process leading to development of a binary predictor, and the use of such a predictor in making Bayesian probability revisions. Here, we refer to data from the example Scenarios *A*, *B* and *D* as described in Madden (59). Readers may also wish to note a similar example placed in a clinical context, where Fischer & Bachmann (25) provide data for three scenarios in which a test based on plasma levels of C-reactive protein is deployed in the diagnosis of sepsis.

The starting point is a normalized prediction-realization table for the FHB data (Table 2). Consider Scenario *A*. From Table 2 we may calculate conditional probabilities representing the properties of the binary predictor as follows: $\Pr(P_1|E_1)$ (sensitivity, *TPP*) = 0.30/0.36 = 0.833, $\Pr(P_2|E_2)$ (specificity, *TNP*) = 0.54/0.64 = 0.844, $\Pr(P_2|E_1)$ (*FNP*) = 1−*TPP* = 0.167, $\Pr(P_1|E_2)$ (*FPP*) = 1−*TNP* = 0.156; the likelihood ratio of a positive prediction is *LR*(+) = 5.333, and the likelihood ratio of a negative prediction is *LR*(−) = 0.198 (Table 3) (all calculations are shown to 3 d.p.). The calculated sensitivity and specificity values are consequent on the choice of a particular operational threshold value. The results of the corresponding calculations for Scenarios *B* and *D* are shown in Tables 2 and 3. We see that in all three scenarios, sensitivity = 0.833 and specificity = 0.844 (so *LR*(+) and *LR*(−) are as above), because all three scenarios are based on the same case and control distributions and the same threshold value. What distinguishes the

scenarios is prior probability: Pr(E_1) = 0.36, 0.05 and 0.85 for scenarios *A*, *B* and *D*, respectively (Table 2). For Scenario *A*, with prior probabilities Pr(E_1) = 0.36 and Pr(E_2) = 0.64, the posterior probabilities are Pr($E_1|P_1$) = 0.750, Pr($E_2|P_1$) = 0.250, Pr($E_2|P_2$) = 0.900, and Pr($E_1|P_2$) = 0.100 (Table 3). The results of the corresponding calculations for Scenarios *B* and *D* are also shown in Table 3.

Now, from an information theoretic perspective, we can consider a prediction as a message that transforms a set of prior probabilities into a corresponding set of posterior probabilities (45). As previously, the prior probabilities are denoted Pr(E_1) and Pr(E_2) for epidemic and non-epidemic, respectively. A message P_i is received which serves to transform these prior probabilities into the posterior probabilities Pr($E_j|P_i$), with $\sum_j \Pr(E_j|P_i) = 1,\ j = 1,2$. Working in base-10 logarithms, the information content of this message with respect to actual status E_j is:

$$information\ content\ of\ P_i = \log_{10}\left(\frac{\Pr(E_j|P_i)}{\Pr(E_j)}\right) \text{Hart}.$$

Of interest is the relationship between information contents and log-likelihood ratios. Here, recalling equation 5, we obtain after some rearrangement:

$$\log_{10}\left(\frac{\Pr(E_1|P_1)}{\Pr(E_1)}\right) - \log_{10}\left(\frac{\Pr(E_2|P_1)}{\Pr(E_2)}\right) = \log_{10}(LR(+))\ \text{Hart} \quad (9)$$

($\log_{10}(LR(+)) > 0$); and similarly recalling equation 7, we obtain:

$$\log_{10}\left(\frac{\Pr(E_1|P_2)}{\Pr(E_1)}\right) - \log_{10}\left(\frac{\Pr(E_2|P_2)}{\Pr(E_2)}\right) = \log_{10}(LR(-))\ \text{Hart} \quad (10)$$

($\log10(LR(-)) < 0$). For further related discussion, see (34).

From Figure 1(*a*), we can see that for P_1 predictions, the information content $\log_{10}(\Pr(E_1|P_1)/\Pr(E_1))$ is large when the prior probability $\Pr(E_1)$ is small, and decreases monotonically towards zero as $\Pr(E_1)$ increases towards one. The largest information gain from a correct P_1 prediction is when $\Pr(E_1)$ is small, the smallest gain is when $\Pr(E_1)$ is large. As the information content $\log_{10}(\Pr(E_1|P_1)/\Pr(E_1))$ becomes decreasingly positive, the information content $\log_{10}(\Pr(E_2|P_1)/\Pr(E_2))$ becomes increasingly negative. Taking Scenarios *A*, *B*, and *D* from Table 3 as examples, we have $\log_{10}(\Pr(E_1|P_1)/\Pr(E_1)) = 0.642, 0.319, 0.056$ Hart and $\log_{10}(\Pr(E_2|P_1)/\Pr(E_2)) = -0.085, -0.408, -0.671$ Hart for $\Pr(E_1) = 0.05$ (Scenario *B*), 0.36 (Scenario *A*), 0.85 (Scenario *D*), respectively. In each case, $\log_{10}(LR(+)) = \log_{10}(\Pr(E_1|P_1)/\Pr(E_1)) - \log_{10}(\Pr(E_2|P_1)/\Pr(E_2))$ (equation 9) = 0.727 Hart.

From Figure 1(*b*), for P_2 predictions, the information content $\log_{10}(\Pr(E_2|P_2)/\Pr(E_2))$ is large when $\Pr(E_2)$ is small and decreases monotonically towards zero as $\Pr(E_2)$ increases towards one. Here, the largest information gain from a correct P_2 prediction is when $\Pr(E_2)$ is small, the smallest gain is when $\Pr(E_2)$ is large. As the information content $\log_{10}(\Pr(E_2|P_2)/\Pr(E_2))$ becomes decreasingly positive, the information content $\log_{10}(\Pr(E_1|P_2)/\Pr(E_1))$ becomes increasingly negative. Again taking the scenarios *A*, *B*, and *D* from Table 3 as examples, we have $\log_{10}(\Pr(E_2|P_2)/\Pr(E_2)) = 0.498, 0.148, 0.018$ Hart and $\log_{10}(\Pr(E_1|P_2)/\Pr(E_1)) = -0.207, -0.556, -0.687$ Hart for $\Pr(E_2) = 0.15$ (*D*), 0.64 (*A*), 0.95 (*B*), respectively. In each case, $\log_{10}(LR(-)) = \log_{10}(\Pr(E_1|P_2)/\Pr(E_1)) - \log_{10}(\Pr(E_2|P_2)/\Pr(E_2))$ (equation 10) = −0.704 Hart.

To summarize: from Figure 1(*a*), for P_1 predictions, the information contents are $\log_{10}(\Pr(E_1|P_1)/\Pr(E_1)) > 0$ and $\log_{10}(\Pr(E_2|P_1)/\Pr(E_2)) < 0$; from Figure 1(*b*), for P_2 predictions, the information contents are $\log_{10}(\Pr(E_2|P_2)/\Pr(E_2)) > 0$ and $\log_{10}(\Pr(E_1|P_2)/\Pr(E_1)) < 0$. We see that the information content of a correct prediction is positive, while the information content of

an incorrect prediction is negative. In relation to Scenarios *A*, *B* and *D*, Madden (59, page 17) notes: "The likelihood ratio is unchanged since this is a property of the predictor, not the prior probability", and this is certainly the case with respect to the numerical values of the likelihood ratios (and so also their logarithms). Like log-likelihood ratios, information contents characterize probability updating from prior to posterior, but unlike log-likelihood ratios, information contents vary with prior probability (Figure 1). Additionally (Figure 1 again), while the numerical values of the log-likelihood ratios do not change, the balance of positive and negative information contents that constitute the log-likelihood ratios changes with prior probability. Thus Figure 1 resolves Langlotz's (56) conundrum by means of a decomposition of log-likelihood ratios, showing that log-likelihood ratios that are numerically equal are not necessarily representative of "the same likelihood ratio" in terms of their constituent information contents.

Weight of Evidence in Practice

A practical approach to the application of log-likelihood ratios in evidence accumulation for clinical diagnosis has been described by Van den Ende et al. (83, 84) Their diagrammatic approach adopts the same base-10 logarithm scale used by Turing's group (an early version of this scale appeared in 51). This diagrammatic approach could have application in crop protection decision making, where there is a natural time sequence for risk accumulation over the growing season. One potential problem to be aware of here is the sequential accumulation of evidence based on risk factors that may be non-independent (83). Hand & Yu (39) provide an interesting discussion of this problem, including reasons why it may not always present an overwhelming difficulty. Still, caution is required; and, if necessary, adjusted weights of evidence can be

calculated to account for non-independence (see (47) for a phytopathological example). The risk accumulation concept was the basis of a new risk assessment method for eyespot disease of winter wheat (a stem-base disease caused by *Oculimacula yallundae* and *O. acuformis*) in the UK (2, 10). The risk assessment is a two-stage process, involving the calculation of pre-sowing risk points and a disease assessment in the spring. Hughes et al. (47) provide further discussion of this kind of two-stage risk assessment.

CONCLUSION

Bayes' rule provides a framework for decision making in disease management (59, 86, 87). The analysis allows current evidence relating to risk factors to be combined with a prior probability of an epidemic, or of the need for a control intervention, resulting in an updated (posterior) conditional probability given the evidence. By stating Bayes' rule in the form of equations 5-8, we can account for evidence relating to risk factors as weights of evidence. Weights of evidence (as $\log_{10}$-likelihood ratios) have some desirable properties in relation to evidence accumulation, both in theory and in practice. Log_{10}-likelihood ratios quantify evidence from risk factors for which data may be recorded on different measurement scales. Evidence is not accumulated additively on a probability scale, but for a decision process based on a series of independent risk factors, the $\log_{10}$-likelihood ratios are additive. This allows a useful graphical interpretation of evidence accumulation (80, 81). And expressing $\log_{10}$-likelihood ratios as weights of evidence provides us with a technical terminology that matches the everyday language interpretation.

We should also be prepared to account for less formal expressions of information, knowledge and data as evidence (36, 37, 38, 76, 85). Collaborative processes, whereby analysts and decision

makers may educate each other about their respective needs and capabilities are the key to risk assessment (26). The evidence that supports risk assessment in disease management decision making comes in a variety of forms, from a variety of sources, and may be used in a variety of ways (46). As seen from a formal analytical perspective, decision support for disease management comprises evidence based on risk factors. But evidence of this type has a wider strategic impact beyond the development and implementation of a system for tactical decision support (e.g., 9, 18, 28, 29, 89). Putting the evidence, rather than the threshold, at the centre of the decision process acknowledges that the framework for decision making in disease management needs to be able to accommodate more than one perspective on the process.

DISCLOSURE STATEMENT

The author has no conflict of interest to declare.

ACKNOWLEDGMENT

SRUC receives grant-in-aid from the Scottish Government.

LITERATURE CITED

1. Aegerter BJ, Nuñez JJ, Davis RM. 2003. Environmental factors affecting rose downy mildew and development of a forecasting model for a nursery production system. *Plant Dis*. 87:732−38

2. AHDB. 2012. *Topic Sheet* 111: *Managing eyespot in winter wheat*. Stoneleigh Park, Warwickshire: Agriculture and Horticulture Development Board. 2 pp.

3. AHDB. 2015. *Information Sheet* 40: *Risk Assessment for Fusarium Mycotoxins in Wheat*. Stoneleigh Park, Warwickshire: Agriculture and Horticulture Development Board. 2 pp.

4. AHDB. 2015. *Oilseed Rape Guide*. Stoneleigh Park, Warwickshire: Agriculture and Horticulture Development Board. 32 pp. 2nd ed.

5. AHDB. 2016. *Barley Disease Management Guide*. Stoneleigh Park, Warwickshire: Agriculture and Horticulture Development Board. 28 pp.

6. AHDB. 2016. *Wheat Disease Management Guide*. Stoneleigh Park, Warwickshire: Agriculture and Horticulture Development Board. 28 pp.

7. Biggerstaff BJ. 2000. Comparing diagnostic tests: a simple graphic using likelihood ratios. *Statist. Med.* 19:649−63

8. Binns MR, Nyrop JP, van der Werf W. 2000. *Sampling and Monitoring in Crop Protection*: *The Theoretical Basis for Designing Practical Decision Guides*. Wallingford, UK: CAB International. 284 pp.

9. Brown SL, Culbreath AK, Todd JW, Gorbet DW, Baldwin JA, Beasley JP. 2005. Development of a method of risk assessment to facilitate integrated management of spotted wilt of peanut. *Plant Dis*. 89:348−56

10. Burnett F, Butler-Ellis C, Hughes G, Knight S, Ray R. 2012. *Forecasting eyespot development and yield losses in winter wheat*. Project Report No. 491, Home-Grown Cereals Authority. 89 pp.

11. Cabinet Office. 2008. *National Risk Register*. London, UK: HMSO. 45 pp.

12. Cabinet Office. 2015. *National Risk Register of Civil Emergencies*. London, UK: HMSO. 57 pp.

13. Caraguel CG, Vanderstichel R. 2013. The two-step Fagan's nomogram: ad hoc interpretation of a diagnostic test result without calculation. *Evid. Based Med.* 18:125−28

14. Carisse O, McRoberts N, Brodeur L. 2008. Comparison of monitoring- and weather-based risk indicators of botrytis leaf blight of onion and determination of action thresholds. *Can. J. Plant Pathol.* 30:442−56

15. Chiarappa L, ed. 1981. *Crop Loss Assessment Methods - Supplement* 3. Slough, UK: Commonwealth Agricultural Bureaux. 123 pp.

16. Collett D. 2003. *Modelling Binary Data*. Boca Raton, FL: Chapman & Hall/CRC. 387 pp. 2nd ed.

17. Cooke BM. 2006. Disease assessment and yield loss. In *The Epidemiology of Plant Diseases*, ed. BM Cooke, DG Jones, B Kaye, pp. 43-80. Dordrecht: Springer. 576 pp. 2nd ed.

18. Culbreath AK, Todd JW, Brown SL. 2003. Epidemiology and management of tomato spotted wilt in peanut. *Annu. Rev. Phytopathol.* 41:53−75

19. De Wolf ED, Madden LV, and Lipps PE. 2003. Risk assessment models for wheat Fusarium head blight epidemics based on within-season weather data. *Phytopathology* 93:428−35

20. Duttweiler KB, Gleason ML, Dixon PM, Sutton TB, McManus PS, Monteiro JEBA. 2008. Adaptation of an apple sooty blotch and flyspeck warning system for the Upper Midwest United States. *Plant Dis*. 92:1215−22

21. Edwards S. 2007. *Investigation of Fusarium mycotoxins in UK wheat production*. Project Report No. 413, Home-Grown Cereals Authority. 86 pp.

22. Edwards S. 2011. *Improving risk assessment to minimise fusarium mycotoxins in harvested wheat grain*. Project Report No. 477, Home-Grown Cereals Authority. 47 pp.

23. Fabre F, Dedryver CA, Leterrier JL, Plantegenest M. 2003. Aphid abundance on cereals in autumn predicts yield losses caused by *Barley yellow dwarf virus*. *Phytopathology* 93:1217−22

24. Fagan TJ. 1975. Nomogram for Bayes's theorem. *N. Engl. J. Med*. 293:257

25. Fischer JE, Bachmann LM, Jaeschke R. 2003. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med*. 29:1043−51

26. Fischhoff B. 2015. The realities of risk-cost-benefit analysis. *Science* 350:DOI: 10.1126/science.aaa6516

27. Fosgate GT. 2005. Letter to the editor. *Statist. Med*. 24:1287–88

28. Gent DH, De Wolf ED, Pethybridge SJ. 2011. Perceptions of risk, risk aversion, and barriers to adoption of decision support systems and integrated pest management: an introduction. *Phytopathology* 101:640−43

29. Gent DH, Mahaffee WF, McRoberts N, Pfender WF. 2013. The use and role of predictive systems in disease management. *Annu. Rev. Phytopathol*. 51:267–89

30. Gent DH, Ocamb CM. 2009. Predicting infection risk of hop by *Pseudoperonspora humuli*. *Phytopathology* 99:1190−98

31. Gent DH, Turechek WW, Mahaffee WF. 2007. Sequential sampling for estimation and classification of the incidence of hop powdery mildew I: leaf sampling. *Plant Dis*. 91:1002−12

32. Gent DH, Turechek WW, Mahaffee WF. 2007. Sequential sampling for estimation and classification of the incidence of hop powdery mildew II: cone sampling. *Plant Dis*. 91:1013−20

33. Giroux M-E, Bourgeois G, Dion Y, Rioux S, Pageau D, Zoghlami S, Parent C, Vachon E, Vanasse A. 2016. Evaluation of forecasting models for Fusarium head blight of wheat under growing conditions of Quebec, Canada. *Plant Dis*. 100:1192−201

34. Good IJ. 1979. Studies in the history of probability and statistics. XXXVII. A. M. Turing's statistical work in World War II. *Biometrika* 66:393−96

35. Good IJ. 1985. Weight of evidence: a brief survey. In: *Bayesian Statistics* 2, ed. Bernardo JM, DeGroot MH, Lindley DV, Smith AFM, pp. 249-70. Amsterdam, The Netherlands: Elsevier Science Publishers BV (North Holland). 778 pp.

36. Gustafson DH, Shukla RK, Delbecq A, Walster GW. 1973. A comparative study of differences in subjective likelihood estimates made by individuals, interacting groups, Delphi groups, and nominal groups. *Organ. Behav. Hum. Perf.* 9:280−91

37. Gustafson LL, Ellis SK, Bartlett CA. 2005. Using expert opinion to identify risk factors important to infectious salmon-anemia (ISA) outbreaks on salmon farms in Maine, USA and New Brunswick, Canada. *Prev. Vet. Med*. 70:17−28

38. Gustafson L, Klotins K, Tomlinson S, Karreman G, Cameron A, Wagner B, Remmenga M, Bruneau N. Scott A. 2010. Combining surveillance and expert evidence of viral hemorrhagic septicemia freedom: a decision science approach. *Prev. Vet. Med.* 94:140–53

39. Hand DJ, Yu K. 2001. Idiot's Bayes - not so stupid after all? *Int. Stat. Rev.* 69:385−98

40. Harikrishnan R, del Río LE. 2008. A logistic regression model for predicting risk of white mold incidence on dry bean in North Dakota. *Plant Dis.* 92:42−46

41. Harremöes P. 2009. *Entropy* – new editor-in-chief and outlook. *Entropy* 11:1−3

42. Hellmich M, Lehmacher W. 2005. A ruler for interpreting diagnostic test results. *Methods Inf. Med.* 44:124−26

43. Herich L, Lehmacher W, Hellmich M. 2015. Drop the likelihood ratio: a novel non-electronic tool for interpreting diagnostic test results. *Methods Inf. Med.* 54:283−87

44. Hughes G. 1999. Sampling for decision making in crop loss assessment and pest management: introduction. *Phytopathology* 89:1080−83

45. Hughes G. 2012. *Applications of Information Theory to Epidemiology*. St. Paul, MN: APS Press. 158 pp.

46. Hughes G, Burnett FJ. 2015. Integrating experience, evidence and expertise in the crop protection decision process. *Plant Dis.* 99:1197−1203

47. Hughes G, Burnett FJ, Havis ND. 2013. Disease risk curves. *Phytopathology* 103:1108−14

48. Hughes G, Gottwald TR. 1998. Survey methods for assessment of citrus tristeza virus incidence. *Phytopathology* 88:715−23

49. Hughes G, Gottwald TR. 1999. Survey methods for assessment of citrus tristeza virus incidence when *Toxoptera citricida* is the predominant vector. *Phytopathology* 89:487−94

50. James WC. 1974. Assessment of plant diseases and losses. *Annu. Rev. Phytopathol.* 12:27−48

51. Jeffreys H. 1939. *Theory of Probability*. Oxford, UK: Clarendon Press. 380 pp.

52. Johnson DA, Alldredge JR, Hamm PB. 1998. Expansion of potato late blight forecasting models for the Columbia Basin of Washington and Oregon. *Plant Dis.* 82:642−45

53. Johnson DA, Alldredge JR. Vackoch DL. 1996. Potato late blight forecasting models for the semi-arid environment of south-central Washington. *Phytopathology* 86:480−84

54. Johnson NP. 2004. Advantages to transforming the receiver operating characteristic (ROC) curve into likelihood ratio co-ordinates. *Statist. Med.* 23:2257−66

55. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes III J. 1961. Factors of risk in the development of coronary heart disease – six-year follow-up experience. *Ann. Intern. Med.* 55:33−50.

56. Langlotz CP. 2003. Fundamental measures of diagnostic examination performance: usefulness for clinical decision making and research. *Radiology* 228:3−9

57. Lindblad M, Waern P. 2002. Correlation of wheat dwarf incidence to winter wheat cultivation practices. *Agr. Ecosyst. Environ.* 92:115–22

58. Looseley ME, Newton AC. 2014. Assessing the consequences of microbial infection in field trials: seen, unseen, beneficial, parasitic and pathogenic. *Agronomy* 4:302−21

59. Madden LV. 2006. Botanical epidemiology: some key advances and its continuing role in disease management. *Eur. J. Plant Pathol*. 115:3−23

60. Madden LV, Hughes G. 1999. Sampling for plant disease incidence. *Phytopathology* 89:1088−103

61. Madden LV, Hughes G, van den Bosch F. 2007. *The Study of Plant Disease Epidemics*. St. Paul, MN: APS Press. 421 pp.

62. McRoberts N, Hughes G, Madden LV. 2003. The theoretical basis and practical application of relationships between different disease intensity measurements in plants. *Ann. appl. Biol.* 142:191−211

63. Mila AL, Driever GF, Morgan DP, Michailides TJ. 2005. Effects of latent infection, temperature, precipitation, and irrigation on panicle and shoot blight of pistachio in California. *Phytopathology* 95:926−32

64. Montgomery DC. 1991. *Introduction to Statistical Quality Control*. New York, NY: John Wiley & Sons. 674 pp. 2nd ed.

65. Nyrop JP, Binns MR, van der Werf W. 1999. Sampling for IPM decision making: where should we invest time and resources? *Phytopathology* 89:1104-11

66. Oxley SJP, Hunter EA. 2005. *Appropriate fungicide doses on winter barley: producing dose-response data for a decision guide*. Project Report No. 366, Home-Grown Cereals Authority. 85 pp.

67. Oxley SJP, Hunter EA. 2009. *Appropriate fungicide doses on winter barley: producing dose-response data for a decision guide*. Project Report No. 455, Home-Grown Cereals Authority. 93 pp.

68. Paul PA, Munkvold GP. 2004. A model-based approach to preplanting risk assessment for gray leaf spot of maize. *Phytopathology* 94:1350−57

69. Paveley ND, Lockley KD, Sylvester-Bradley R, Thomas J. 1997. Determinants of fungicide spray decisions for wheat. *Pestic. Sci.* 49:379−88

70. Pethybridge SJ, Gent DH, Esker PD, Turechek WW, Hay FS, Nutter FW Jr. 2009. Site-specific risk factors for ray blight in Tasmanian pyrethrum fields. *Plant Dis.* 93:229‑37

71. Pethybridge SJ, Gent DH, Hay FS. 2011. Epidemics of ray blight on pyrethrum are linked to seed contamination and overwintering inoculum of *Phoma ligulicola* var. *inoxydabilis*. *Phytopathology* 101:1112−21

72. Plant RE. 1986. Uncertainty and the economic threshold. *J. Econ. Entomol.* 79:1−6

73. Prandini A, Sigolo S, Filippi L, Battilani P, Piva G. 2009. Review of predictive models for Fusarium head blight and related mycotoxin contamination in wheat. *Food Chem. Toxicol.* 47:927−31

74. Seem RC. 1984. Plant disease incidence and severity relationships. *Annu. Rev. Phytopathol.* 22:137−50

75. Selvin S. 1996. *Statistical Analysis of Epidemiologic Data*. New York, NY: Oxford University Press, Inc. 510 pp. 3rd ed.

76. Sherman J, Gent DH. 2014. Concepts of sustainablility, motivations for pest management approaches, and implications for communicating change. *Plant Dis.* 98:1024−35

77. Stern VM, Smith RF, van den Bosch R, Hagen KS. 1959. The integrated control concept. *Hilgardia* 29:81−101

78. Swets JA. 1973. The relative operating characteristic curve in psychology. *Science* 182:990−1000

79. Swets JA, Dawes RM. Monahan J. 2000. Better decisions through science. *Sci. Am.* 283:70−75

80. Turechek WW, Ellis MA, Madden LV. 2001. Sequential sampling for incidence of Phomopsis leaf blight of strawberry. *Phytopathology* 91:336−47

81. Turechek WW, Wilcox WF. 2005. Evaluating predictors of apple scab with receiver operating characteristic curve analysis. *Phytopathology* 95:679−91

82. Twengström E, Sigvald R, Svensson C, Yuen J. 1998. Forecasting Sclerotinia stem rot in spring sown oilseed rape. *Crop Prot*. 17:405−11

83. Van den Ende J, Bisoffi Z, Van Puymbroek H, Van Gompel A, Derese A, Lynen L, Moreira J, Janssen PAJ. 2007. Bridging the gap between clinical practice and diagnostic clinical epidemiology: pilot experiences with a didactic model based on a logarithmic scale. *J. Eval. Clin. Pract*. 13:374–80

84. Van den Ende J, Moreira J, Basinga P, Bisoffi Z. 2005. The trouble with likelihood ratios. *Lancet* 366:548

85. VHSV Expert Panel and Working Group. 2010. Viral hemorrhagic septicemia virus (VHSV IVb) risk factors and association measures derived by expert panel. *Prev. Vet. Med*. 94:128–39

86. Yuen JE, Hughes G. 2002. Bayesian analysis of plant disease prediction. *Plant Pathol*. 51:407−12

87. Yuen J, Mila A. 2004. Are Bayesian approaches useful in plant pathology? In *Bayesian Statistics and Quality Modeling in the Agro-Food Production Chain* Vol. 3, ed. van Boekel MAJS, Stein A, van Bruggen AHC, pp. 95-103. Dordrecht, The Netherlands: Kluwer Academic Publishers. 165 pp.

88. Yuen J, Twengström E. Sigvald R. 1996. Calibration and verification of risk algorithms using logistic regression. *Eur. J. Plant Pathol.* 102:847−54

89. Zadoks JC. 1985. On the conceptual basis of crop loss assessment: the threshold theory. *Annu. Rev. Phytopathol.* 23:455−73

90. Zweig MH, Campbell G. 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.* 39:561−77

Table 1 The generic prediction-realization table for a predictor with two categories of actual status E_j, j = 1,2; and two categories of predicted status P_i, i = 1,2. [a]

	Realization, E_j		
Prediction, P_i	E_1	E_2	Row sums
P_1	$\Pr(P_1 \cap E_1)$	$\Pr(P_1 \cap E_2)$	$\Pr(P_1)$
P_2	$\Pr(P_2 \cap E_1)$	$\Pr(P_2 \cap E_2)$	$\Pr(P_2)$
Column sums	$\Pr(E_1)$	$\Pr(E_2)$	1

[a] The table shows:

$\Pr(E_1)$: Prior probability of requirement for a control intervention.

$\Pr(E_2)$: Prior probability of no requirement for a control intervention.

$\Pr(P_1)$: Probability of a prediction of need for a control intervention.

$\Pr(P_2)$: Probability of a prediction of no need for a control intervention.

$\Pr(P_i \cap E_j)$: In the body of the table are the joint probabilities.

Table 2 Estimated probabilities for the prediction-realization tables corresponding to Scenarios *A*, *B* and *D* from (59).

Probability [a]	Scenario *A*	Scenario *B*	Scenario *D*
$\Pr(P_1 \cap E_1)$	0.300	0.042	0.708
$\Pr(P_1 \cap E_2)$	0.100	0.148	0.023
$\Pr(P_2 \cap E_1)$	0.060	0.008	0.142
$\Pr(P_2 \cap E_2)$	0.540	0.802	0.127
$\Pr(E_1)$	0.360	0.050	0.850
$\Pr(E_2)$	0.640	0.950	0.150
$\Pr(P_1)$	0.400	0.190	0.732
$\Pr(P_2)$	0.600	0.810	0.268

[a] See Table 1.

Table 3 Estimated quantities corresponding to Scenarios *A*, *B* and *D* from (59).

Quantity	Scenario *A*	Scenario *B*	Scenario *D*
$\Pr(P_1\|E_1)$ [a]	0.833	0.833	0.833
$\Pr(P_1\|E_2)$ [a]	0.156	0.156	0.156
$\Pr(P_2\|E_2)$ [a]	0.844	0.844	0.844
$\Pr(P_2\|E_1)$ [a]	0.167	0.167	0.167
$LR(+)$ [b]	5.333	5.333	5.333
$LR(-)$ [b]	0.198	0.198	0.198
$\Pr(E_1\|P_1)$ [c]	0.750	0.219	0.968
$\Pr(E_1\|P_2)$ [c]	0.100	0.010	0.528
$\Pr(E_2\|P_2)$ [c]	0.900	0.990	0.472
$\Pr(E_2\|P_1)$ [c]	0.250	0.781	0.032

[a] $\Pr(P_i|E_j) = \Pr(P_i \cap E_j)/\Pr(E_j)$ (see Table 2).

[b] $LR(+) = \Pr(P_1|E_1)/\Pr(P_1|E_2)$; $LR(-) = \Pr(P_2|E_1)/\Pr(P_2|E_2)$.

[c] $\Pr(E_j|P_i) = \Pr(P_i \cap E_j)/\Pr(P_i)$ (see Table 2).

Figure 1

Information contents of log-likelihood ratios (denominated in hartleys). Calculations are based on sensitivity = 0.833, specificity = 0.844. Information contents calculated from Tables 2 and 3 are indicated by points marked ○, with the prior probabilities of the three scenarios indicated by points marked ● on the horizontal axis. $\Pr(E_2) = 1-\Pr(E_1)$). (*a*) P_1 predictions. The short-dashed horizontal line shows $\log_{10}(LR(+)) = 0.727$ Hart. The solid curve shows information content $\log_{10}(\Pr(E_1|P_1)/\Pr(E_1))$; the long-dashed curve shows information content $\log_{10}(\Pr(E_2|P_1)/\Pr(E_2))$. At any value of $\Pr(E_1)$, $\log_{10}(\Pr(E_1|P_1)/\Pr(E_1)) - \log_{10}(\Pr(E_2|P_1)/\Pr(E_2)) = \log_{10}(LR(+)) = 0.727$ Hart. (*b*) P_2 predictions. The short-dashed horizontal line shows $\log_{10}(LR(-)) = -0.704$ Hart. The solid curve shows information content $\log_{10}(\Pr(E_1|P_2)/\Pr(E_1))$; the long-dashed curve shows information content $\log_{10}(\Pr(E_2|P_2)/\Pr(E_2))$. At any value of $\Pr(E_2)$, $\log_{10}(\Pr(E_1|P_2)/\Pr(E_1)) - \log_{10}(\Pr(E_2|P_2)/\Pr(E_2)) = \log_{10}(LR(-)) = -0.704$ Hart.