Plant Pathology (2002) **51**, 407–412

LETTER TO THE EDITOR

Bayesian analysis of plant disease prediction

J. E. Yuen[a]*† and G. Hughes[b]

[a]*Department of Ecology and Crop Production, Swedish University of Agricultural Sciences, SE 750 05 Uppsala, Sweden; and*
[b] *Institute of Ecology and Resource Management, University of Edinburgh, Edinburgh EH9 3JG, UK*

Rule-based systems for the prediction of the occurrence of disease can be evaluated in a number of different ways. One way is to examine the probability of disease occurrence before and after using the predictor. Bayes's Theorem can be a useful tool to examine how a disease forecast (either positive or negative) affects the probability of occurrence, and simple analyses can be conducted without knowing the risk preferences of the targeted decision makers. Likelihood ratios can be calculated from the sensitivity and specificity of the forecast, and provide convenient summaries of the forecast performance. They can also be used in a simpler form of Bayes's Theorem. For diseases where little or no prior information on occurrence is available, most forecasts will be useful in that they will increase or decrease the probability of disease occurrence. For extremely common or extremely rare diseases, likelihood ratios may not be sufficiently large or small to substantially affect the probability of disease occurrence or make any difference to the actions taken by the decision maker.

Introduction

We often seek better methods to manage plant diseases and pests. BLITECAST (Krause *et al*., 1975) was one of the first such computerized systems, designed for the control of potato late blight (caused by *Phythophthora infestans*), though the first decision making aid was proposed for apple scab (caused by *Venturia inaequalis*) (Mills, 1994) and predated modern computer systems. Often they are referred to as expert systems (Travis & Latin, 1991) although plant disease simulation models can be used in a variety of different applications, as learning or decision tools, in a strategic or tactical context (Teng & Yuen, 1991). In this sense, the tactical decision tools, such as EPIPRE (Zadoks, 1981) are those designed to be used for decision-making throughout the growing season.

A common feature is that all of these systems, computerized or not, utilize a rule-based approach to crop protection decision-making. In the simplest such examples, it is the occurrence or otherwise of the disease or pest that is predicted. More sophisticated examples predict the level of disease or of the pathogen or pest population, in relation to some threshold, above which crop protection measures will be applied, either once or several times through the growing season.

Like all predictors, these rule-based methods sometimes provide incorrect predictions. The frequency of incorrect predictions will influence the rate of adoption of a method by potential users, and its continued use. This letter seeks to illustrate the relationship between the performance (here used to refer to the frequency of correct predictions) of a method and disease occurrence, using Bayes's theorem.

*To whom correspondence should be addressed.

†E-mail: Jonathan.Yuen@evp.slu.se

Accepted 23 April 2002

Statistical methods: background and examples

An example with eyespot

Consider first an example, taken from Jones (1994). In a study of the use of fungicides to control eyespot disease of wheat (caused by *Pseudocercosporella herpotrichoides*), data collected during the 1980s from 58 sites in England were presented. Crops were classified by their yield response to prochloraz treatment at GS30-31; the treatment was either justified (yield response $\geq 0{\cdot}2$ t ha^{-1}) or not (yield response $< 0{\cdot}2$ t ha^{-1}). The predictor was the incidence of eyespot at GS 30–31; at $\geq 20\%$ tillers affected the need to apply treatment was indicated, at $< 20\%$ tillers affected the indication was to withhold treatment. The results were given in a 2×2 table (see Table 1).

The sensitivity of a predictor expresses the frequency of positive predictions (in this example, predictions indicating the need to apply treatment) as a proportion of the total number that were actually positive (in this case, those where the treatment was justified). So here, sensitivity = 28/41 = 0·68 (Table 1). The specificity of a predictor expresses the frequency of negative predictions (in this case, predictions where the indication was to

Table 1 Eyespot predictor

		predictor		
		apply treatment	withhold treatment	total
True status	treatment justified	28	13	41
	treatment not justified	10	7	17

withhold treatment) as a proportion of the total number that were actually negative (in this case, those where the treatment was unjustified). So here, specificity = 7/17 = 0·41 (Table 1).

The sensitivity of a predictor is also referred to as the true positive rate, and the specificity as the true negative rate. From Table 1 we can also characterize the false negative rate (= 13/41 = 0·32, i.e. 1 – sensitivity) and the false positive rate (= 10/17 = 0·59, i.e. 1 – specificity).

We now consider the general problem, and introduce some notation. Let A represent disease occurrence (omitting references to thresholds, for simplicity). $\Pr(A)$ is the unconditional probability of disease occurrence. In Bayesian terms, $\Pr(A)$ represents the prior probability of disease based on our past experience. There is also a prior probability of nonoccurrence, denoted $\Pr(\bar{A})$, such that $\Pr(A) + \Pr(\bar{A}) = 1$. Now, let B represent a prediction of the occurrence of disease and $\bar{B}$ a prediction of the nonoccurrence of disease. Sensitivity (the true positive rate) is the conditional probability $\Pr(B|A)$ (read as 'the probability of a prediction of disease occurrence given that disease actually occurred'). Similarly, specificity (the true negative rate) is the conditional probability $\Pr(\bar{B}|\bar{A})$. The false negative rate is $\Pr(\bar{B}|A)$ and the false positive rate is $\Pr(B|\bar{A})$. These conditional probabilities characterize properties of the predictor: the extent to which the occurrence or nonoccurrence of disease is indicated, in the knowledge either that disease actually did occur, or actually did not.

From the point of view of practical predictions about the need for crop protection measures, we are usually less interested in sensitivity and specificity as such, and more concerned with conditional probabilities such as $\Pr(A|B)$, the probability of disease occurrence, given a prediction of occurrence; and $\Pr(\bar{A}|\bar{B})$, the probability of nonoccurrence of disease, given a prediction of nonoccurrence. $\Pr(A|B)$ is sometimes referred to as the positive predictive value (denoted PPV), and $\Pr(\bar{A}|\bar{B})$ as the negative predictive value (denoted NPV). In Bayesian terms, $\Pr(A|B)$ and $\Pr(\bar{A}|\bar{B})$ represent the posterior probability of the event given the data (i.e. of the actual occurrence of disease given the prediction of occurrence or the actual nonoccurrence of disease given a prediction of nonoccurrence, respectively).

$\Pr(A|B)$ can be related to $\Pr(B|A)$ using Bayes's theorem as follows:

$$\Pr(A|B) = \frac{\Pr(B|A)\ \Pr(A)}{\Pr(B|A)\ \Pr(A) + \Pr(B|\bar{A})\ \Pr(\bar{A})} \qquad (1)$$

A similar relationship can be derived for the relationship of $\Pr(\bar{A}|\bar{B})$ to $\Pr(\bar{B}|\bar{A})$:

$$\Pr(\bar{A}|\bar{B}) = \frac{\Pr(\bar{B}|\bar{A})\ \Pr(\bar{A})}{\Pr(\bar{B}|\bar{A})\ \Pr(\bar{A}) + \Pr(\bar{B}|A)\ \Pr(A)} \qquad (2)$$

Thus the positive and negative predictive values can be calculated from the sensitivity and the specificity of the predictor and the unconditional probability (prior probability) of disease occurrence.

One useful measure that contains information from both specificity and sensitivity is the likelihood ratio of a positive prediction, denoted $\mathrm{LR}(A,B)$:

$$\mathrm{LR}(A,B) = \frac{\Pr(B|A)}{1 - \Pr(\bar{B}|\bar{A})} \qquad (3)$$

i.e. $\mathrm{LR}(A,B)$ = sensitivity/(1– specificity). In this example, $\mathrm{LR}(A,B) = 0{\cdot}68/0{\cdot}59 = 1{\cdot}15$.

Likewise, we can also define the likelihood ratio of a negative prediction, denoted $\mathrm{LR}(A,\bar{B})$:

$$\mathrm{LR}(A,\bar{B}) = \frac{1 - \Pr(B|A)}{\Pr(\bar{B}|\bar{A})} \qquad (4)$$

i.e. $\mathrm{LR}(A,\bar{B})$ = (1– sensitivity)/specificity. In this example, $\mathrm{LR}(A,\bar{B}) = 0{\cdot}32/0{\cdot}41 = 0{\cdot}78$.

The likelihood ratios of positive and negative predictions are also useful in that they yield a simplified version of Bayes's theorem. If we rewrite the prior probabilities as odds ($\mathrm{odds}(event) = \Pr(event)/(1 - \Pr(event))$), we can write instead of eqn 1:

$$\mathrm{odds}(A|B) = \mathrm{odds}(A) \times \mathrm{LR}(A,B) \qquad (5)$$

and also derive similar relationship for the probability of disease after a negative prediction

$$\mathrm{odds}(A|\bar{B}) = \mathrm{odds}(A) \times \mathrm{LR}(A,\bar{B}) \qquad (6)$$

In these expressions of Bayes's theorem, the posterior odds is equal to the prior odds multiplied by a likelihood ratio (Sackett *et al.*, 1995).

From equation 5, we can see that as long as $\mathrm{LR}(A,B) > 1$, the effect of a prediction of disease occurrence is to increase the posterior odds of disease occurrence relative to the prior odds. From equation 6, we can see that as long as $\mathrm{LR}(A,\bar{B}) < 1$, the effect of a prediction of disease nonoccurrence is to decrease the posterior odds of disease occurrence relative to the prior odds. Both of these attributes are desirable in a good predictor of disease. To summarize, from the point of view of making predictions of disease occurrence, we would ideally like a predictor to have $\mathrm{LR}(A,B)$ as large as possible, and simultaneously have $\mathrm{LR}(A,\bar{B})$ as small as possible (Biggerstaff, 2000).

In order to use Bayes's theorem, we need information

Table 2 Prior probabilities and odds for the occurrence of Sclerotinia stem rot in Uppland and Västmanland, based on 20 years averages, or single worst year

Region	time period	Pr(A)	odds(A)
Uppland	20 year average	0·16	0·19
Uppland	single year	0·61	1·56
Västmanland	20 year average	0·23	0·30
Västmanland	single year	0·64	1·78

Table 3 Sensitivity and specificity of varying decision thresholds, and likelihood ratios resulting from positive and negative predictions for the occurrence of Sclerotinia stem rot

Threshold risk points score	sensitivity	specificity	LR(A,B)	LR(A,$\bar{B}$)
35	0·90	0·77	3·913	0·130
40	0·77	0·84	4·812	0·274
50	0·35	0·95	7·000	0·684

on the prior probabilities. For the purpose of illustration, we return to the eyespot example and use data from Cook *et al.* (1991) for winter wheat crops in England and Wales between 1985 and 1989. Over this period, the average percentage of crops untreated with prochloraz and affected by moderate to severe eyespot varied between 11·1% and 18·9%. The weighted average over six years, taking into account 633 crops, was 16·4%, so we take Pr(A) = 0·16 or odds(A) = 0·19. Then, using the likelihood ratios calculated above, from equation 5, odds($A|B$) = 0·23 and from equation 6, odds($A|\bar{B}$) = 0·15. After a prediction of disease occurrence the odds of disease occurrence increases, and after a prediction of disease nonoccurrence the odds of disease occurrence decreases. This is as we would wish, but the changes in odds of disease occurrence after a prediction are not large.

An example with Sclerotinia stem rot

Sclerotinia stem rot of oil seed rape (caused by *Sclerotinia sclerotiorum*) can be a problem in Sweden in some years (Twengström, 1999). Surveys of fields in Uppland and Västmanland were started in 1979, where unsprayed fields were classified into different disease incidence classes. In addition, some farmers in these regions who applied fungicides were asked to leave a 50 m swath without fungicides in order to to see the effects of the fungicides, and these unsprayed areas were also classified in a similar manner. A 20 year average (1979–98) for the frequency of need for fungicide application in Uppland and Västmanland, as well as the highest probabilities in a single year (E. Twengström, personal communication, 2000, Table 2), were used as prior probabilities of the need for fungicide application for this example. Earlier records did not indicate how many fields were surveyed, but for the latter half of the period (from 1987 to 1998) a total of 923 fields from both regions were surveyed. Fields were considered to need a fungicide application if the disease incidence was greater than 25%.

In the case of Sclerotinia stem rot, predictions of the need for fungicide application are made by means of an algorithm (calibrated in risk points) that combines information on cropping history, previous occurrence of the disease and weather (Twengström *et al.*, 1998). Risk factors are classified into different categories, assigned different point values, which are then summed to give the risk points. For this example, information on the sensitivity and specificity of the prediction algorithm was presented in the form of a receiver operating characteristic (ROC) curve by Twengström *et al.* (1998). An ROC curve is a graphical plot of the true positive rate as a function of the false positive rate at all possible decision thresholds of the prediction algorithm (Metz, 1978; Yuen *et al.*, 1996). Thus, for any given implementation of the prediction algorithm, sensitivity and specificity can be calculated from the plot.

In practice, the appropriate risk point threshold for the prediction algorithm published by Twengström *et al.* (1998) will vary, depending (among other things) on the decision maker's attitude to risk. In this example, thresholds for fungicide application of 35, 40 and 50 risk points were used (Twengström *et al.*, 1998). The various sensitivity and specificity combinations obtained are shown in Table 3. A threshold of 40 points results in approximately equal false positive and false negative rates. Increasing the threshold to 50 points reduces the false positive rate, but increases the false negative rate. Reducing the threshold to 35 points reduces the false negative rate but increases the false positive rate. The corresponding values of the likelihood ratios LR(A,B) and LR(A,$\bar{B}$) can be calculated from equations 3 and 4. The resulting values of LR(A,B) and LR(A,$\bar{B}$) are shown in Table 3. None of the thresholds discussed here is superior overall. A threshold of 50 is the best for confirming disease occurrence (highest LR(A,B)), but the worst for confirming nonoccurrence (highest LR(A,$\bar{B}$)). A threshold of 35 is the best for confirming the nonoccurrence of disease (lowest LR(A,$\bar{B}$)) but the worst for confirming disease occurrence (lowest LR(A,B)). A threshold of 40 risk points is intermediate between the others from both points of view.

The values of LR(A,B) and LR(A,$\bar{B}$) can be combined with data in Table 2, via Bayes's theorem (eqns 5 and 6), in order to characterize the posterior odds, given the prediction. These values are shown in Table 4. At any value of the prior odds of disease occurrence, a higher LR(A,B) gives a higher posterior odds of disease occurrence, given a prediction of disease; and a predictor with a lower LR(A,$\bar{B}$) gives a lower posterior odds of disease occurrence, given a prediction of disease nonoccurrence. Adopting a threshold risk point score of 50 leads to the largest increases in posterior odds of disease after a prediction of disease occurrence. Adopting a threshold risk point score of 35 leads to the largest decreases in posterior odds of disease after a prediction of disease nonoccurrence. Again, we note that the magnitude of these effects depends on the prior odds of disease occurrence.

Table 4 Increases and decreases in the probability of Sclerotinia occurrence after positive and negative prediction based on varying decision thresholds. Values in body of table are odds, probabilities are shown in brackets

Threshold risk points score	LR(A,B) or LR(A,$\bar{B}$)	*odds*(A) after prediction			
		0·19	1·56	0·30	1·78
35	3·913	0·75 (0·43)	6·12 (0·86)	1·17 (0·54)	6·96 (0·87)
40	4·812	0·92 (0·48)	7·53 (0·88)	1·44 (0·59)	8·56 (0·90)
50	7·000	1·33 (0·57)	10·95 (0·92)	2·09 (0·68)	12·44 (0·93)
35	0·130	0·02 (0·024)	0·20 (0·17)	0·04 (0·037)	0·23 (0·19)
40	0·274	0·05 (0·050)	0·43 (0·30)	0·08 (0·076)	0·49 (0·33)
50	0·684	0·13 (0·12)	1·07 (0·52)	0·20 (0·17)	1·22 (0·55)

Discussion

The performance of predictive rules or algorithms can be summarized by the sensitivity and specificity, or as likelihood ratios. If an algorithm had sensitivity and specificity of 0·9 it would have values for LR(A,B) and LR(A,$\bar{B}$) of 9 and 0·11, respectively. If, before a prediction, the probability of disease occurrence is 0·1, a positive prediction increases the chance of disease occurrence to 0·5, while a negative prediction decreases the chance to 0·01. Now consider the former result in particular. A positive prediction has increased the chance of disease occurrence, by a factor of five, but the probability of disease occurrence is still only 0·5: disease occurrence and nonoccurrence are equally likely, despite the positive prediction. If disease is infrequent, even this reasonably good predictor does not raise the posterior probability of disease occurrence above 0·5. A prediction of the nonoccurrence of disease decreases the posterior probability of disease occurrence to 0·01, but this may not be regarded as of much help (since disease was infrequent anyway).

Consider the case of frequent disease. With a prior probability of disease of 0·9, a positive prediction then increases the chance of disease occurrence to 0·99, while a negative prediction decreases the chance to 0·5. A negative prediction has decreased the chance of disease occurrence, but the probability of disease occurrence is still 0·5: disease occurrence and nonoccurrence are equally likely, despite the negative prediction. A prediction of the occurrence of disease increases the posterior probability of disease occurrence, but this may not be regarded as of much help (since disease was frequent anyway).

We now turn to the Sclerotinia example. The results for Uppland and Västmanland are qualitatively similar, so we just discuss the former. An additional cautionary note is that there is some overlap in the data used to develop the predictor and the data used to calculate the long-term averages, but a separation of the data would have been an extremely tedious task, and as presented here, the example is intended to be illustrative rather than prescriptive. There are two main differences between the scenario for this example and for the hypothetical example discussed above. First, for Sclerotinia the prior probability of disease lies between the extremes of the frequent case (Pr(A) = 0·9) and the infrequent case (Pr(A) = 0·1) for the previous hypothetical example. If the 20 years average as the low end of the range and the single year value as the high end are taken, the relevant values are Pr(A) = 0·16 and Pr(A) = 0·61, respectively (Table 2). Second, the best predictor for confirming disease occurrence (threshold risk score = 50) has a lower LR(A,B) than the hypothetical predictor in the previous example, and the best predictor for confirming the nonoccurrence of disease (threshold risk score = 35) has a higher LR(A,$\bar{B}$) than the hypothetical predictor.

Table 4 has the relevant results written in terms of posterior odds and probabilities of disease occurrence. Using the best predictor for confirming the occurrence of disease, a positive prediction increases the prediction of disease occurrence from 0·16 (using the 20 years average prior) to 0·57, or from 0·61 to 0·92 (using the single year prior). A negative prediction decreases the prediction of disease occurrence from 0·16 (using the 20 years average prior) to 0·12, or from 0·61 to 0·52 (using the single year prior). Using the best predictor for confirming the nonoccurrence of disease, a positive prediction increases the prediction of disease occurrence from 0·16 (using the 20 years average prior) to 0·43, or from 0·61 to 0·86 (using the single year prior). A negative prediction decreases the prediction of disease occurrence from 0·16 (using the 20 years average prior) to 0·024, or from 0·61 to 0·17 (using the single year prior).

We note in particular that even though the best predictor for confirming disease occurrence (threshold risk score = 50) has a lower LR(A,B) than the hypothetical predictor in the previous example, use of the predictor with a threshold risk score equal to 50 increases the prediction of disease occurrence, given a positive prediction from 0·16 (using the 20 years average prior) to 0·57. In the previous example, use of the hypothetical predictor increased the prediction of infrequent disease, given a positive prediction, from 0·1 to 0·5. A more helpful prediction is obtained for Sclerotinia because its prior is not so low.

Generally, it is impractical to develop prediction systems for diseases that fall in to the infrequent or frequent occurrence categories. For infrequently occurring diseases, the prior probability of disease is low, so a test with an extremely high LR(A,B) would be required in order for the test to increase the posterior probability of disease

to a point where action might be taken. For frequently occurring diseases, LR(A,$\bar{B}$) for a test would have to be very low in order to reduce the posterior probability of disease occurrence to a level at which no action would be taken. Wider acceptance of predictors might be expected for diseases that are neither infrequent nor frequent. Even predictors with modest performance might be useful in this situation. Note, however, that the results for the eyespot example show an example where disease is neither infrequent nor frequent, but where the performance of the putative predictor is poor. Here, Pr(A) is taken to be 0·16, but relatively low LR(A,B) and relatively high LR(A,$\bar{B}$) mean that the posterior probability of disease given a positive prediction increases only to 0·18, and the posterior probability of disease given a negative prediction decreases only to 0·13. Eyespot assessment early in the season is known to be an unreliable indicator of subsequent disease development (Scott & Hollins, 1978). The results presented here perhaps reflect this more clearly than a statement of overall accuracy (cost effectiveness of treatment accurately predicted at 60% of the 58 sites; Jones, 1994) based on the data shown in Table 1.

The prediction of the occurrence of Sclerotinia stem rot in oil seed rape has been well accepted by advisors and farmers, and the simple rules based on the studies by Twengström *et al.* (1998), have been used in a number of forms, both in paper and via the Internet, http://www.tvs.slu.se or http://www.ipm.evp.slu.se. The wider acceptance of this predictive system might be due to a number of different factors. The system has relatively high sensitivity and specificity, and varying decision thresholds can be implemented to allow for differences in risk attitudes. In addition, the prior probability of disease occurrence is neither extremely high nor extremely low. While it is difficult to know what the exact prior probability is, since the numbers used here assume no time change in the occurrence of the disease, even after the worst years, the prior probability is only slightly greater than 50% if we use the single worst year alone as the prior.

In reality, we have often little information on prior occurrence of pests or on the performance of the predictors. Even the use of historical data in order to estimate the prior probabilities involves assumptions about unchanging weather, cultivars, cultural practices, and pathogen populations. These assumptions may or may not be true. In addition, small data sets may carry high variability due to a limited number of observations. Few predictive systems have addressed the issue of sensitivity and specificity, although the general issue of uncertainty has been discussed in other systems (Wiles *et al.*, 1993; Gold & Sutton, 1986; Gold & Sutton, 1989). Despite the relative ease with which these performance characteristics can be calculated, few studies have actually presented the sensitivity and specificity of prediction methods.

The use of Bayes's theorem in disease prediction and control was first proposed by Carlson (1969, 1970), within the context of payoff. Thus, the prior and conditional probabilities were those related to payoffs or utility, not pest occurrence. Utility is a dimensionless real number that reflects user's preference for risky situations (Anderson *et al.*, 1977; Gold, 1989; Teng & Yuen, 1991). Derivation of the probabilities related to utility, which in this case are subjective, often takes place via questions directed at decision makers (Carlson, 1969).

Even in the presence of reasonably sound priors and information about sensitivity and specificity, our Bayesian calculations and logic remain a theoretical discussion when it comes to predicting human behaviour, and acceptance or rejection of decision support systems. What matters is not the true prior (that is established via historical data, surveys, or other objective means) but the perceived prior of the decision maker. How well this coincides with an objective measure has not been studied, but one could guess that these subjective priors are weighted. For example, an individual might give more weight to recent observations and less weight to those that occurred some time in the past. In the Sclerotinia example, this discrepancy can be represented by the 20 year average, which gives equal weight to all observations and the single highest value, which gives all the weight to the worst year and no weight to the other years.

Likewise, it is not the true LRs that result from a positive or negative prediction, but those that are perceived by the decision maker. If he mistrusts the system, he may ascribe lower sensitivity and/or specificity to the system, thus changing the likelihood ratio. Overconfidence and enthusiastic developers may lead to higher LRs than are really justified. Deriving subjective probabilities related to payoff or utility (Carlson, 1969) would eliminate this problem, but would require much more than the analysis of historical data, and lead to methods that are much more complex than those presented here. One advantage to a subjective probability approach is that the decision maker's reactions to different risky situations will be included.

Still, an objective study of the performance of the pest prediction systems combined with estimations of pest occurrence will lead to a better understanding of the potential usefulness that these systems have. These objective analyses can be conducted without knowing the relevant utility functions. Knowing the performance of a system will enable the targeting of areas where there is a chance that it might be used. In the absence of a system, knowledge about pest occurrence can be used to determine the minimum performance criteria necessary for a predictive system to succeed.

Acknowledgements

This work was supported in part by a grant from the Swedish Council for Forestry and Agricultural Research.

References

Anderson JR, Dillon JL, Hardaker HB, 1977. *Agricultural Decision Analysis*. Ames, IA, USA: Iowa State University Press.

Biggerstaff BJ, 2000. Comparing diagnostic tests: a simple graphic using likelihood ratios. *Statistics in Medicine* **19**, 649–63.

Carlson GA, 1969. Bayesian analysis of pesticide use. *Proceedings of the American Statistic Association. Business and Economic Statistics Section*, 411–416.

Carlson GA, 1970. A decision theoretic approach to crop disease prediction and control. *American Journal of Agricultural Economics* **52**, 216–23.

Cook RJ, Polley RW, Thomas MR, 1991. Disease-induced losses in winter-wheat in England and Wales 1985–89. *Crop Protection* **10**, 504–8.

Gold HJ, 1989. Decision analytic modeling for plant disease control. In: Leonard, K, Fry, W, eds. *Plant Disease Epidmiology*, Vol. 2 *Genetics, Resistance and Management*. New York, NY, USA: McGraw-Hill, 84–122.

Gold HJ, Sutton TB, 1986. A decision analytic model for chemical control of sooty blotch and flyspeck diseases of apple. *Agricultural Systems* **21**, 129–57.

Jones DR, 1994. Evaluation of fungicides for control of eyespot disease and yield loss relationships in winter wheat. *Plant Pathology* **43**, 831–46.

Krause RA, Massie LB, Hyre RA, 1975. BLITECAST: a computerized forecast of potato late blight. *Plant Disease Reporter* **59**, 95–8.

Metz CE, 1978. Basic principles of ROC analysis. *Seminars in Nuclear Medicine* **8**, 283–98.

Mills WD, 1944. *Efficient Use of Sulfur Dusts and Sprays During Rain to Control Apple Scab*. Ithaca, NY, USA: New York State Agricultural Experimental Station, Cornell Extension Bulletin 630.

Sackett DL, Haynes RB, Tugwell P, 1985. *Clinical Epidemiology – a Basic Science for Clinical Medicine*. Boston, MA, USA: Little, Brown & Co.

Scott PR, Hollins TW, 1978. Prediction of yield loss due to eyespot in winter wheat. *Plant Pathology* **27**, 125–31.

Teng PS, Yuen JE, 1991. Epidemic models: lessons from plant pathology. In: Levin, M, Strauss, H, eds. *Risk Assessment in Genetic Engineering*. New York, NY, USA: McGraw-Hill, 272–96.

Travis JW, Latin RX, 1991. Development, implementation, and adoption of expert systems in plant pathology. In: Cook, R, Zentmeyer, G, Cowling, E, eds. *Annual Review of Phytopathology*, **29**, 343–60.

Twengström E, 1999. *Epidemiology and Forecasting of Sclerotinia Stem Rot on Spring Sown Oilseed Rape in Sweden*. Uppsala, Sweden: Swedish University of Agricultural Sciences, PhD Thesis.

Twengström E, Sigvald R, Svensson C, Yuen JE, 1998. Forecasting sclerotinia stem rot in spring sown oilseed rape. *Crop Protection* **17**, 405–11.

Wiles LG, Gold HJ, Wilkerson GG, 1993. Modeling the uncertainty of weed density estimates to improve post-emergence herbicide control decisions. *Weed Research* **33**, 241–52.

Yuen JE, Twengström E, Sigvald R, 1996. Calibration and verification of risk algorithms using logistic regression. *European Journal of Plant Pathology* **102**, 847–54.

Zadoks JC, 1981. EPIPRE: a disease and pest management system for winter wheat. *EPPO Bulletin* **11**, 365–9.