European Journal of Plant Pathology (2006) 115: 95–103
DOI 10.1007/s10658-005-1230-z

Disease assessment concepts and the advancements made in improving the accuracy and precision of plant disease data

Forrest W. Nutter Jr.[1] Paul D. Esker[1] and Rosalee A. Coelho Netto[2]
[1]*Department of Plant Pathology, Iowa State University, 351 Bessey Hall, Ames, IA, 50011 USA (Fax: +1-515-294-9420; E-mail: fwn@iastate.edu);* [2]*National Research Institute of the Amazon, Caixa Postal 478, CEP 69011-670, Manaus, AM, Brazil*

Accepted 13 July 2005

Key words: phytopathometry

Abstract

New concepts in phytopathometry continue to emerge, such as the evolution of the concept of pathogen intensity versus the well-established concept of disease intensity. The concept of pathogen severity, defined as the quantitative measurement of the amount of pathogen per sampling unit has also emerged in response to the now commonplace development of quantitative molecular detection tools. Although the concept of disease severity, i.e., the amount of disease per sampling unit, is a well-established concept, the accuracy and precision of visual estimates of disease severity is often questioned. This article will review disease assessment concepts, as well as the methods and assessment aides currently available to improve the accuracy and precision of visually-based disease severity data. The accuracy and precision of visual disease severity assessments can be improved by quantitatively measuring and comparing the accuracy and precision of rates and/or assessment methods using linear regression, by using computer-based disease assessment training programmes, and by developing and using diagrammatic keys (standard area diagrams).

Introduction

"How can plant pathologists apply advanced statistical procedures or develop quantitative predictive models based upon disease assessment data of unknown accuracy and precision?" David R. Mackenzie, 1979

The efficient application of any integrated disease management programme requires accurate and precise information concerning the quantitative measurement of the disease and/or pathogen population, yet the accuracy and precision of quantitative disease/pathogen assessments in plant pathology is often taken for granted (Main, 1977; Zadoks and Schein, 1979; Gaunt, 1987; Kranz, 1988; Nutter et al., 1991; Nutter and Schultz, 1995; Nutter and Gaunt, 1996). Accurate and precise disease (or pathogen) assessments provide a quantitative link between disease management theory and practice (Shokes et al., 1987; Nutter et al., 1991; Nutter and Schultz, 1995). An integral component of studying the interactions of host and pathogen populations in time and space is the ability to accurately discriminate between levels of injury (disease intensity) caused by plant pathogens.

Disease intensity is a generic term for the amount of disease in a host population. Disease intensity can be either the independent variable or the dependent variable in stimulus–response models; however, in both cases, disease intensity needs to be quantified with a high degree of accuracy and precision if meaningful predictive models are to be developed (Nutter, 1990; O'Brien and van Bruggen, 1992; Nutter and Gaunt, 1996; Guan and Nutter, 2003). For example, quantifying disease intensity–crop yield (loss) relationships demands a high degree of accuracy and precision

with regards to disease assessments because disease intensity is used as the independent variable (X) in single-point or area under the disease progress curve (AUDPC) – crop yield (loss) regression models (Guan and Nutter, 2001, 2004). The stimulus (disease intensity) must be measured accurately in order to develop yield response or yield loss models that have adequate predictive capabilities.

The symbol Y is often used to represent a measure of disease (or pathogen) intensity because disease intensity assessments are often graphed on the y-axis with respect to time (X-axis). The graphical representation of disease intensity versus time is referred to as a disease progress curve, whereas the graphical representation of pathogen intensity versus time is referred to as a pathogen progress curve (Nutter, 2001). A disease (or pathogen) progress curve is the signature of an epidemic and represents the integration of all host, pathogen, and environmental effects (including pathogen vectors and disease management tactics) that occur during the period of host–pathogen interaction (Campbell and Madden, 1990; Nutter, 1997b).

Quantification of disease intensity also requires a high degree of accuracy and precision when disease intensity is the dependent variable (Y) to quantify the rate of disease progress with respect to time (X), or the change in disease intensity with respect to distance (X). In both cases, X (time or distance) can be measured with great accuracy, and therefore, the accuracy and precision of disease intensity assessments (Y) directly affect how much of the variation in Y can be explained by X in such models.

Disease assessment defined

The branch of plant pathology that deals with the theory and practice of quantitative disease (and/or pathogen) assessment is known as phytopathometry (Main, 1977; Zadoks and Schein, 1979; Campbell and Madden, 1990). Disease assessment is defined as the act (or process) of quantitatively measuring disease intensity (Campbell and Madden, 1990; Nutter et al., 1991; Nutter and Gaunt, 1996). In plant pathology, there are two basic and distinct populations that can be quantitatively assessed: the pathogen population and the disease population (Nutter, 1997b, 1999). Because plant disease epidemics result from the interaction of host and pathogen populations in time and space, as affected by the environment, quantification of the disease population usually involves an assessment of visible injury (disease symptoms). This is true because disease injury is often directly proportional to the size of the pathogen population (Nutter et al., 1991; Nutter and Guan, 2001). On the other hand, pathogen assessments can be obtained by directly measuring the pathogen population (e.g. the number of spores, sclerotia, nematodes, etc.) per unit area or volume, or the use of a detection method to determine the presence or absence of a pathogen for each sampling unit (e.g., ELISA or PCR to detect the presence of a pathogen in or on a sampling unit). Thus, researchers can perform disease assessments or pathogen assessments (or both); however, these terms should not be used interchangeably because they represent different populations being assessed (Nutter, 1997b, 1999, 2001).

Disease versus pathogen intensity

Disease intensity (Y) is a general (generic) term used for quantifying the amount of disease in a population (Campbell and Madden, 1990; Nutter et al., 1991; Nutter and Gaunt, 1996). In plant pathology, the three most common measures of disease (and pathogen) intensity are: (i) prevalence, (ii) incidence, and (iii) disease severity. Disease prevalence is a term that is often used interchangeably (and mistakenly) with disease incidence. Prevalence is defined as the number of geographical sampling units (fields, farms, counties, states, regions, etc.) where a disease or pathogen has been detected, divided by the total number of geographical sampling units assessed (Zadoks and Schein, 1979; Campbell and Madden, 1990; Nutter et al., 1991; Nutter and Gaunt, 1996). It is important to distinguish disease prevalence from pathogen prevalence. Disease prevalence measures the proportion (or percentage) of geographical sampling units (fields, counties, countries, etc.) where a disease (expressing symptoms) has been found to occur, divided by the total number of geographical sampling units inspected or surveyed, whereas pathogen prevalence is a measure of the number of geographical sampling units where the pathogen has been detected (e.g., by direct plating, inspections for the presence of

pathogen signs, ELISA, PCR, etc.), divided by the total number of geographical sampling units that were inspected, tested, or indexed (Nutter, 2001). Prevalence data are often multiplied by 100 to express as a percentage. A single diseased or infected plant (or plant part) is all that is required to change the status of a geographic sampling unit from negative (−) to positive (+), provided the sensitivity of the method is sufficient to detect the presence of a pathogen in a bulked (diluted) sample. Bulking samples is particularly useful when pathogen incidence is low because the number of bulked samples tested or indexed is often less than the number of individuals sampled and processed, thus reducing the cost of detection per sampling unit (Nutter and Gaunt, 1996; Nutter, 1997b).

Disease incidence is defined as the number of sampling units (e.g., leaflets, leaves, stems, tillers, whole plants, seeds, etc.) that are diseased (expressing symptoms), divided by the total number of sampling units sampled and assessed (Nutter et al., 1991; Nutter, 1997b, 1999). As with prevalence, it is important to make a clear distinction as to whether incidence is based on detection of the pathogen or on the basis of disease (visual symptoms) in a host population (Nutter, 2001). Progress curves based on pathogen detection (indexing) methods, such as ELISA, may closely mirror progress curves based on disease symptoms (Padgett et al., 1990; Nutter 2001); however, in many instances, the use of different disease assessment or pathogen detection methods may result in progress curves with quite different shapes and rates (Nutter, 1997b, 2001).

Disease severity is a measure of the amount of disease per sampling unit and it is this particular type of measurement that this article will focus upon (Nutter et al., 1991; Nutter, 1997b). Researchers should clearly define disease severity by providing not only a descriptive definition, but also an operational definition that includes the dimensions that were used to assess disease. In plant pathology, disease severity is most often operationally defined as the diseased leaf area (l^2), divided by the total leaf area of a leaf or sampling unit (L^2)$\times$100, i.e., ($l^2/L^2\times100$) to obtain percentage disease severity (James, 1971; Nutter et al., 1991). Other common measures of disease severity include the number of lesions/leaf (or sampling unit), the number of lesions/cm^2 of leaf (or other sampling unit), or the area of non-green tissue of a sampling unit divided by the total area of the sampling unit $\times$ 100 (Nutter, 2001). Disease severity could be also defined as the volume of a disease-induced gall (using the equation and dimensions for a cylinder or a sphere, etc.), as is done in human epidemiology for cancerous tumors (Nutter, 1999).

The concept of pathogen severity is becoming more widespread as new methods are developed to quantify the amount of a pathogen present in a sampling unit. Examples include the use of quantitative PCR methods to estimate the amount of virus (or pathogen) present per gram of leaf tissue, or the number of nematodes per gram of root tissue. In spite of advances concerning the concept of pathogen severity, the majority of severity assessments employed today usually involve visual estimates of disease severity ($l^2/L^2\times100$), and yet, a number of critical questions still remain. At the top of the list of such questions is: how can we do better to improve the accuracy and precision of visual disease assessments? This article will address specific ways to improve the accuracy and precision of diseases assessments. These are: (i) use of regression to quantify the precision of disease raters and/or assessment methods, (ii) the use of computer-based assessment training programmes, and (iii) the use of standard area diagrams (diagrammatic keys) in colour.

Use of linear regression to assess and compare the precision of assessment methods

The expenditure of time and money to develop, evaluate, and compare disease assessment methods can prevent serious flaws (e.g., rater bias) in data acquisition. Disease assessment methods should provide accurate and precise information that satisfies the goals and needs of the research (Nutter and Gaunt, 1996). Campbell and Madden (1990) have defined precision as the lack of variation in disease estimates when the same sampling units are evaluated by other raters. However, this definition of precision excludes another potential source of error, i.e., the repeatability of individual raters (Nutter et al., 1993). Shokes et al. (1987) proposed using a test-retest procedure using correlation analysis to quantify rates repeatability; however, this method provides a measure of precision

(agreement) among raters and does not quantify the degree of bias among raters.

Simple linear regression provides a powerful method to quantify the degree of error (bias) due to raters or assessment methods (Nutter et al., 1993). Regression analysis has been used to determine the relative precision of a visual assessment method (disease severity) and a remote sensing assessment method (reflectance at 600 nm) in which a hand-held, multispectral radiometer was employed (Nutter et al., 1993). The disease assessed was dollar spot of bentgrass (caused by *Sclerotinia homoeocarpa*). The precision of different disease assessment methods and raters can be evaluated and compared by operationally defining intra-rater repeatability and inter-rater reliability (Nutter and Schultz, 1995). Intra-rater repeatability for different assessment methods can be determined by regressing one set of measurements (Y) (obtained by each rater) with a repeated set of measurements (X) performed on the same set of sampling units. The parameters and statistics used to compare the intra-rater repeatability of different assessment methods and/or raters are: slope, intercept, coefficient of determination (R^2), coefficient of variation (CV), and the standard error of the estimate for Y (SEEy) (Nutter et al., 1993; Nutter and Schultz, 1995). A slope significantly less than or greater than 1.0 would indicate the presence of systematic bias and the greater the deviation from 1.0, the greater the systematic bias for a specific rater and/or method. This is because for each 1% increase in estimated disease severity the first time a set of sampling units is assessed, there should be a corresponding 1% increase in estimated disease severity when the same set of sampling units are assessed a second time by the same rater or method (Nutter et al., 1993). An intercept significantly different from zero indicates the presence of another form of bias that is constant for all disease levels evaluated. The use of R^2, CV, and SEEy values to quantify and compare the precision of disease assessment methods or raters has been previously described (Nutter et al., 1993; Nutter and Schultz, 1995; Nutter, 2001). Likewise, linear regression can be used to quantify precision among raters or methods (inter-rater reliability) by having two or more raters (and/or methods) assess the same set of sampling units, and then evaluating the slopes, intercepts, R^2, CV, and SEEy values (Nutter et al., 1993).

Disease assessment training with computer programmes

The accuracy and precision of disease severity assessments have come into question due to the measurable bias that different raters have shown when evaluating the same set of diseased sampling units (Sherwood et al., 1983; Forbes and Jeger, 1987; Kranz, 1988; Nutter et al., 1993). Accuracy can be defined as the measure of the closeness of a disease assessment to the true value (Nutter et al., 1991; Zadoks and Schein, 1979). When assessing disease severity, the stimulus (X) is the actual disease severity of a sampling unit and the rater's estimate of disease severity (Y) is the response. For each 1% increase in actual severity, we would expect a rater to also to estimate a 1% increase, i.e., the slope should be equal to 1.0 (no systematic bias) and the intercept should not be significantly different from zero (no constant bias present). Accuracy cannot be properly evaluated unless the researcher is confident that the actual (true) disease severity can be measured absolutely. This is easily achieved using computer-generated images of diseased leaves because the computer can be programmed to calculate the number of non-green (diseased) pixels in an image, divided by the total number of pixels in the image $\times$ 100 to obtain a true measure of percentage disease severity (Nutter and Litwiller, 1998; Nutter et al., 2000). The use of computer programmes to enhance learning has become widely accepted for several reasons (Nutter, 1997a). One advantage of computer-aided disease assessment training is that a full range of disease severity levels can be presented as stimuli to which operators of the programme respond. Nutter and Worawitlikit (1990) built upon the computer-based disease assessment training concept by developing a computer programme to assess diseases of peanut called Disease.Pro. Recognizing the tremendous potential to improve the accuracy and precision of disease assessments through computer-based training programmes, Nutter and Litwiller (1998) later developed a more generic disease assessment training programme (Severity.Pro) that allowed the user to select from a menu of leaf shapes (alfalfa, apple, barley, cucumber, grape, tomato, etc.) and lesion types (anthracnose, blotch, downy mildew, target spot, powdery mildew, etc.) to mimic almost any foliar pathosystem. Severity.Pro was recently rewritten

in Java to be more compatible with present-day operating systems.

The most current version of Severity.Pro allows raters to: (i) choose whether or not they want the actual severity to be immediately displayed (feedback), (ii) choose the number of leaves to be assessed and the size of the lesions that will appear on diseased leaf images (small, medium, large, or random), (iii) view graphs of the absolute error (Y) versus the actual severity (X) (absolute error is defined as the estimated severity minus the corresponding actual severity), and (iv) regress the rater's severity estimates (Y), against the actual disease severities (X). These changes allowed for a more powerful training tool because: (i) raters can take a pre-test (without feedback before training) to provide a baseline of how different disease severity levels are perceived, (ii) the data can be viewed in graphical form and analyzed by regression, (iii) raters can execute a drill and practice session and receive feedback as to the actual level of disease severity immediately after the estimated severity is keyed in, (iv) rater improvement and, more importantly, the degree of rater improvement can be documented by having raters take a post-test (without feedback after training) and then compare pre- and post-test regression parameters and statistics, and (v) the results of pre- and post-tests can also be used to evaluate and compare rater performance.

Computer-based disease assessment training programmes provide a useful platform for teaching disease assessment theory and hands-on practice. For example, the results of computer training for six raters using Severity.Pro are shown in Table 1. The six raters evaluated computer-generated images of grapevine leaves infected by downy mildew by assessing 30 images before and after training. In pre-tests, one rater (Rater 2) generally overestimated downy mildew severity throughout the range of the severities tested, with rater error being as high as 21% (solid circles, Figure 1a). Following training, Rater 2's estimates were within 5–10% of the actual severity levels (open circles, Figure 1a). Figure 1b shows that this rater also had a constant bias of 6.8% prior to training (Y-intercept) and that this bias was reduced to near zero (−0.14%) after computer-based training using Severity.Pro. Based on R^2 values, the precision of Rater 2 was also significantly improved following training (R^2 was 95% following training compared with 85% prior to training). As a group, five of the six raters demonstrated improvement in precision following computer training, as measured by improvement in the coefficients of determination (R^2) (Table 1). Three of the raters had R^2 values that were 6–11% higher after training. The other three raters (rater 1, rater 4, and rater 6) were already highly precise and their R^2 values increased 2, 1, and 0%, respectively.

Table 1. Y-intercepts[a], slopes[b], coefficients of determination (R^2)[c], and standard errors of the Y-estimate (SEE_Y)[d] for six raters before training (pretest) and after training (posttest) using a computer programme that simulates downy mildew of grapevines (adapted from Nutter, 2001)

Rater	Pretest				Posttest			
	Intercept	Slope	R^2	SEE_Y	Intercept	Slope	R^2	SEE_Y
1	−7.19	1.11	0.94	2.74	−1.52	1.01	0.96	1.95
2[e]	6.83	1.02	0.85	3.80	−0.14	0.94	0.95	1.97
3	−6.30	0.91	0.91	2.83	7.21	0.83	0.97	1.41
4	1.89	1.06	0.91	2.97	0.61	0.82	0.92	2.25
5	−1.13	1.27	0.84	4.69	−8.46	1.05	0.94	2.61
6	1.77	1.01	0.97	1.62	−3.10	1.03	0.97	1.76
Improved					3/6	2/6	5/6	5/6

[a] Y-intercepts that deviate from zero indicate the presence of a constant source of rater bias with regards to accuracy.
[b] Slopes that deviate from 1.0 indicate the presence of a systematic source of rater bias with regards to accuracy.
[c] The higher the coefficient of determination (R^2), the higher the precision of rater estimates.
[d] The lower the standard error of the Y-estimate, the higher the precision of rater estimates.
[e] Data for Rater 2 are shown in graphical form in Figure 1 as this data would appear in the disease assessment computerized training programme Severity.Pro (Nutter and Litwiller, 1998).

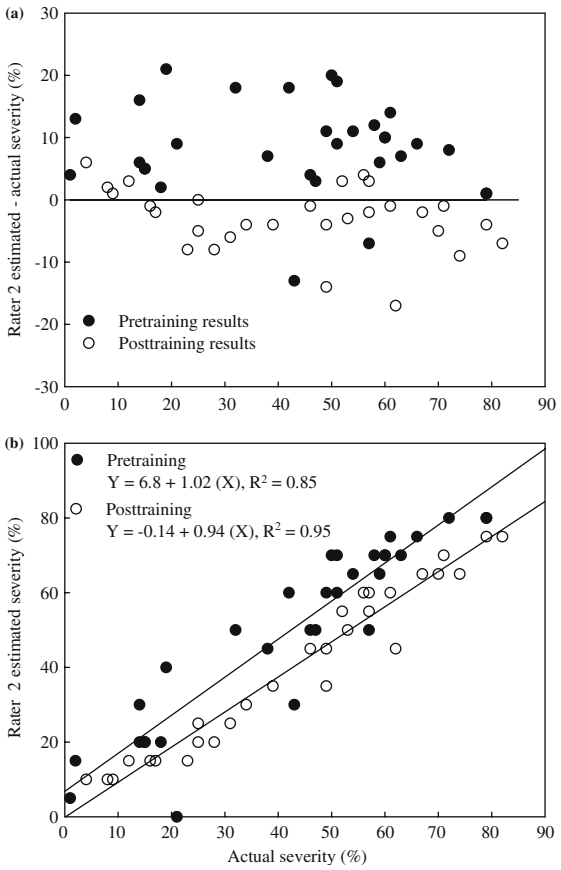

Figure 1. Improvement in the (a) absolute error (estimated minus actual disease severity) and (b) accuracy (slope, intercept) and precision (R^2, SEEy) of Rater 2 (from Table 1), before and after disease assessment using the computer program Severity.Pro (Nutter and Litwiller, 1998).

As stated earlier, the standard error of the *Y*-estimate is another important measure of rater precision. This statistic provides information concerning the degree of error associated with a predicted value of *Y*. Therefore, the lower the SEE_Y, the higher the precision (Nutter and Schultz, 1995). For five of the six raters, SEE_Y values decreased following disease assessment training by an average of 40% (3.41 to 2.04). Rater 6 showed no significant change in pre-test versus post-test SEE_Y values because this rater was already very precise (i.e., rater 6 had very low SEE_Y values in both pre- and post-training tests). Thus, computerized disease assessment training programmes provide an important educational tool that can be used to teach disease assessment theory and concepts, as well as to substantially improve both the accuracy and the precision of disease assessment data.

Use of standard area diagrams to improve the accuracy and precision of disease severity assessments

Disease assessment keys, also known as diagrammatic keys or standard area diagrams are pictorial diagrams that depict the true amount of injury (usually disease severity) on individual sampling units (quadrats, whole plants, leaves, fruit, tubers, etc.). Disease severity of each individual diagram is expressed as a percentage of the total surface area of each sampling unit (disease area/total area of the image $\times$ 100) (Nutter and Esker, 2001). Standard area diagrams (SADs) provide raters with a series of reference images that are accepted to be the truth in terms of the actual amount of injury (severity) depicted on each disease diagram. Clive James developed and marketed the first series of black and white standard area diagrams (James, 1971). More recently, Nutter and Litwiller (1998) developed and tested a computer programme (Severity.Pro) that generates standard area diagrams in colour. Thus, Severity.Pro provides a powerful tool to generate, capture and print diseased leaf images with known severity levels in colour (Nutter et al., 1998). This enables researchers to create a series of pictorial colour diagrams that can be used as an assessment aid to improve the accuracy and precision of disease assessment data. Although it has long been assumed that the use of standard area-diagrammatic keys will help to improve the accuracy and precision of visual disease severity assessments performed by raters (James, 1971; Horsfall and Cowling, 1978; Kranz, 1988), only recently have definitive studies been conducted to demonstrate that the accuracy and precision of disease assessments are actually improved when standard area-diagrammatic keys are used (Godoy et al., 1997; Nutter et al., 1998; Leite and Amorim, 2002; Gomes et al., 2004). As part of a class exercise for students enrolled in a course in plant disease epidemiology at Iowa State University, 10 raters were asked to assess 30 diseased leaf images (representing a range of disease severities) of downy mildew of grape, both with and without the use of colour-standard area diagrams (Nutter and Litwiller, 1998). When individual rater estimates

were regressed against the true disease severity levels as calculated by the computer programme, it was found that rater estimates of disease severity were much closer to the actual (true) severity levels when raters used standard area diagrams as an assessment aid to assess disease severity (Nutter and Esker, 2001).

As an example, Figure 2 shows a typical situation regarding accuracy and precision of visual assessments performed by one rater with, and without, the use of standard area diagrams for grapevine downy mildew. When using the standard area diagrams, this rater had greater accuracy (less systematic and constant bias) as indicated by a slope closer to 1.0 (1.05) and a *Y*-intercept closer to zero (−0.28%) compared to the slope (0.86) and intercept (5.34%) when not using the standard area diagrams. Moreover, R^2 values were higher and SEE_y values were lower when standard area diagrams were used, indicating there was a significant increase in the precision of the assessment data when using the standard area diagrams.

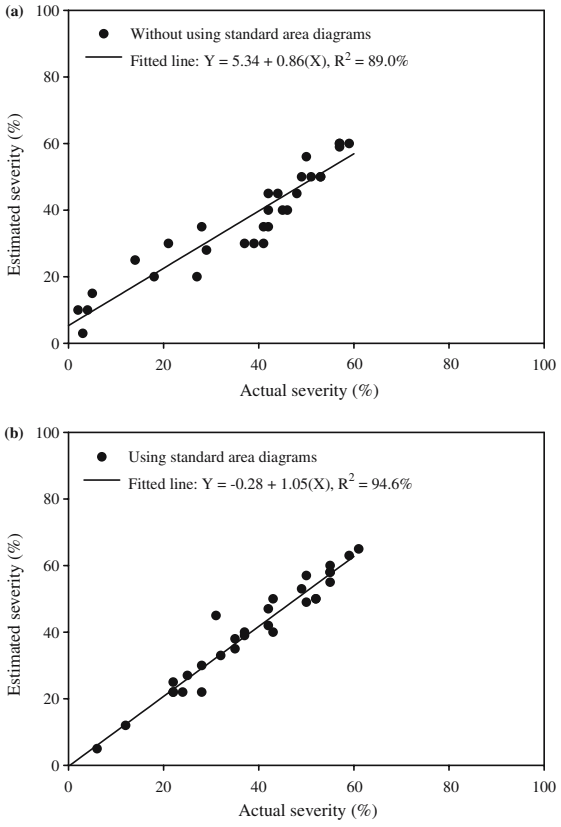

Figure 2. (a) Estimated severity of grapevine downy mildew compared with actual (true) severity when assessing computer images without the use of standard area diagrams and (b) Estimated versus actual severity when using standard area diagrams. Improvements were apparent as both systematic and constant bias were reduced.

As a class, statistical analyses for accuracy showed that eight of the ten raters achieved intercepts closer to zero (less constant bias) and that seven of the ten raters achieved slopes closer to 1.0 (less systematic bias) when standard area diagrams were used (Nutter and Schultz, 1995). Statistical analyses for precision showed that seven of the ten raters achieved higher coefficients of determination (R^2), and eight of the ten raters had lower standard errors of the estimate for *Y* (SEE_y) when standard area-diagrammatic keys were used as an assessment aid. Thus diagrammatic standard area-assessment keys can substantially improve both the accuracy and the precision of visual disease assessments.

Summary and conclusions

The potential for rater bias (under- or over-estimation of the actual level of disease severity) is an ever-present concern that should receive serious consideration by researchers when raters are making visual disease assessments and will use that information as the basis to develop stimulus–response models, or to evaluate and compare disease management tactics, strategies, or integrated disease management systems (Zadoks and Schein, 1979; Gaunt, 1995; Nutter, 1997b, 1999, 2001). Rater bias, however, can be effectively reduced. Disease assessment training programmes using computer-generated images of disease leaves have been shown to improve both accuracy and precision (Nutter and Schultz, 1995; Nutter and Parker, 1997; Nutter and Litwiller, 1998). Moreover, studies by Godoy et al. (1997), Gomes et al. (2004), Nutter (2001), and Nutter and Esker (2001) have documented that the use of standard area diagrams as an assessment aid for visually assessing disease severity can also significantly improve the accuracy and precision of disease severity assessment data. The use of both computer-based disease assessment training programmes and standard area diagrams to improve the accuracy and precision of disease assessment data are not mutually exclusive, as both methods should be used to obtain the best disease assessment data possible. Finally, the use of

regression to evaluate and compare the accuracy and precision of: (i) colour disease assessment protocols (e.g., use of a linear scale versus a logarithmic disease assessment scale) (Nutter and Esker, 2005), (ii) disease assessment instruments (e.g., image analysis or remote sensing sensors and instruments) (Nutter, 1990; Guan and Nutter, 2004), and/or (iii) disease raters, can provide researchers with statistical methods to determine the accuracy and precision of disease assessment data (Nutter et al., 1993; Nutter and Littrell, 1996; Nutter, 1997a; Guan and Nutter, 2003). Thus, researchers can and should place greater focus upon evaluating, comparing, and selecting the best disease assessment protocols, instruments, and/or raters that best meet the goals of the research.

References

Campbell CL and Madden LV (1990) Introduction to Plant Disease Epidemiology, John Wiley, New York.

Forbes GA and Jeger MJ (1987) Factors affecting the estimation of disease intensity in simulated plant structures. Zeitschrift fur Pflanzen Krankheiten und Pflanzenschutz 94: 113–120.

Gaunt RE (1995) New philosophies and technologies in disease measurement and appraisal. Canadian Journal of Plant Pathology 17: 185–189.

Godoy CV, Carneiro MTPG, Iamauti MT, Dalla Pria M, Amorim L, Berger RD and Bergamim Filho A (1997) Diagrammatic scales for bean diseases: development and validation. Journal of Plant Diseases and Protection 104: 336–345.

Gomes AMA, Michereff SJ and Mariano RLR (2004) Development and validation of a diagrammatic key for Cercospora leaf spot of lettuce. Summa Phytopathologica 30: 38–42.

Guan J and Nutter FW, Jr. (2001) Factors affecting the quality and quantity of sunlight reflected from alfalfa canopies. Plant Disease 85: 865–874.

Guan J and Nutter FW, Jr. (2003) Quantifying the intra-rater repeatability and inter-rater reliability of visual disease and remote sensing assessment methods in the alfalfa foliar disease pathosystem. Canadian Journal of Plant Pathology 25: 143–149.

Guan J and Nutter FW, Jr. (2004) Comparison of single-point alfalfa yield models based on visual disease intensity and remote sensing assessments. Canadian Journal of Plant Pathology 26: 314–324.

Gaunt RE (1987) Measurement of disease and pathogens. In: Teng PS (ed.) Crop Loss Assessment and Pest Management (pp. 6–18) The American Phytopathological Society, St. Paul, MN.

Horsfall JG and Cowling EB (1978) Pathometry: the measurement of plant disease. In: Horsfall JG and Cowling EB (eds) Plant Disease: An Advanced Treatise. Vol. 2 (pp. 120–136) Academic Press, New York.

James C (1971) A Manual of Assessment Keys for Plant Disease. Canadian Department of Agriculture Publishing.

Kranz J (1988) Measuring plant disease. In: Kranz J and Rotem J (eds) Experimental Techniques in Plant Disease Epidemiology (pp. 35–50) Springer-Verlag, New York.

Leite RMVBC and Amorim L (2002) Development and validation of a diagrammatic scale for Alternaria leaf spot of sunflower. Summa Phytopathologica 28: 14–19.

Main CE (1977) Crop destruction – The raison d' être of plant pathology. In: Horsfall JG and Cowling EB (eds) Plant Disease: An Advanced Treatise (pp. 55–78) Academic Press, New York.

Nutter FW, Jr. (1990) Remote sensing and image analysis for crop loss assessment. In: Crop Loss Assessment in Rice (pp. 93–105) International Rice Research Institute, Manila, Philippines.

Nutter FW, Jr. (1997a) Disease severity assessment training. In: Francl L and Neher D (eds) Exercises in Plant Disease Epidemiology (pp. 1–7) APS Press, St. Paul, MN.

Nutter FW, Jr. (1997b) Quantifying the temporal dynamics of plant virus epidemics: a review. Crop Protection 16: 603–618.

Nutter FW, Jr. (1999) Understanding the interrelationship between botanical, human, and veterinary epidemiology: The Ys and Rs of it all. Ecosystem Health 5: 131–140.

Nutter FW, Jr. (2001) Disease assessment terms and concepts. In: Maloy OC and Murray TD (eds) Encyclopedia of Plant Pathology (pp. 312–323) John Wiley and Sons, Inc, NY.

Nutter FW, Jr. and Esker PD (2006) Disease assessment keys. In: Maloy OC and Murray TD (eds) Encyclopedia of Plant Pathology (pp. 323–326) John Wiley and Sons, Inc, NY.

Nutter FW, Jr. and Esker PD (2005) The role of psychophysics in phytopathology: the Weber-Fechner law revisited. European Journal of Plant Pathology 114: 199–213.

Nutter FW, Jr. and Gaunt RE (1996) Recent developments in methods for assessing disease losses in forage/pasture crop. In: Chakraborty S, Leath DT, Skipp RA, Pederson GA, Bray RA, Latch GCM and Nutter FW Jr. (eds) Pasture and Forage Crop Pathology (pp. 93–118) American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Madison, WI.

Nutter FW, Jr., Gleason ML, Jenco JH and Christians NC (1993) Assessing the accuracy, intra-rater repeatability and inter-rater reliability of disease assessment systems. Phytopathology 83: 806–812.

Nutter FW, Jr. and Guan J (2001) Disease losses. In: Maloy OC and Murray TD (eds) Encyclopedia of Plant Pathology (pp. 340–351) John Wiley and Sons, Inc, NY.

Nutter FW, Jr. and Littrell RH (1996) Relationships among percent defoliation, canopy reflectance, and pod yield in the peanut-late leafspot system. Crop Protection 15: 135–142.

Nutter FW, Jr. and Litwiller D (1998) A computer program to generate standard area diagrams to aid raters in assessing disease severity. Phytopathology 88: S117.

Nutter FW, Jr., Litwiller D and Wegulo SN (1998). Do standard area diagrams improve the accuracy and precision of disease assessments? Proceedings of the 7th International Congress of Plant Pathology. p. 2.1.16.

Nutter FW, Jr. and Parker SK (1997) Fitting disease progress curves using EPIMODEL. In: Francl L and Neher D (eds) Exercises in Plant Disease Epidemiology (pp. 24–28) APS Press, St. Paul, MN.

Nutter FW, Jr. and Schultz PM (1995) Improving the accuracy and precision of disease assessment: selection of methods and use of computer-aided training programs. Canadian Journal of Plant Pathology 17: 174–184.

Nutter FW, Jr., Teng PS and Shokes RM (1991) Disease assessment terms and concepts. Plant Disease 75: 1187–1188.

Nutter FW, Jr., Wegulo SN, Litwiller D, Emmitt RW and Magarey PA (2000) A computer program to improve the accuracy and precision of severity assessments in grapevines. In: Magarey, PA, Thiele, SA, Tschirpig, KL, Emmet, RW, Clark, K and Magarey RD (eds) Proceedings of the Third International Workshop on Grapevine Downy and Powdery Mildew (pp. 64–69). South Australian Research and Development Institute Series Report No. 50.

Nutter FW, Jr. and Worawitlikit O (1990) Disease. Pro: A computerized disease assessment training program. Proceedings of the American Peanut Research and Ed. Society 22: 38.

O'Brien RD and van Bruggen AHC (1992) Accuracy, precision, and correlation to yield loss of disease severity scales for corky root of lettuce. Phytopathology 82: 91–96.

Padgett GB, Nutter FW, Jr. and Kuhn CW (1990) Quantification of disease resistance that reduces the rate of tobacco etch virus epidemics in bell pepper. Phytopathology 80: 451–455.

Sherwood RT, Berg CC, Hoover MR and Zeiders KE (1983) Illusions in visual assessment of Stagonospora leaf spot of orchardgrass. Phytopathology 73: 173–177.

Shokes FM, Berger RD, Smith DH and Rasp JM (1987) Reliability of disease assessment procedures: A case study with late leafspot of peanut. Oleagineux 42: 245–251.

Zadoks JC and Schein RD (1979) Epidemiology and Plant Disease Management, Oxford University Press, New York.