Eur J Plant Pathol (2014) 139:413–422
DOI 10.1007/s10658-014-0400-2

Development and validation of a standard area diagram set to aid assessment of severity of loquat scab on fruit

Elisa González-Domínguez · Ricardo Brainer Martins · Emerson Medeiros Del Ponte · Sami Jorge Michereff · José García-Jiménez · Josep Armengol

Accepted: 10 February 2014 / Published online: 21 February 2014

Abstract A standard area diagram set (SAD) to aid visual assessment of loquat scab (caused by *Fusicladium eriobotryae*) severity on fruit was developed and evaluated for improving accuracy, precision and reliability of visual estimates. The SAD set contains eight black and white diagrams of diseased fruit with severity values from 2 % to 98 %. To evaluate the SADs, a group of 20 raters (comprising 10 'experienced' and 10 'inexperienced' raters) assessed the same set of 50 images three times, the first without SADs and the second and third using the SADs as an aid. Only for the group of inexperienced raters did SADs significantly improve accuracy (bias correction factor, C_b=0.93 without SADs and 0.98 with SADs), precision (correlation coefficient, r=0.88 without SADs and r=0.96 with SADs) and overall agreement (Lin's concordance correlation coefficient, ρ_c=0.82 without SADs and ρc with SADs = 0.95) of the estimates. Accuracy and precision of the estimates by inexperienced raters were significantly higher than those obtained by the experienced raters, especially for the second assessment with SADs. Inter-rater reliability was improved when SADs were used by inexperienced raters, whereas a high degree of intra-rater reliability was obtained by both experienced and inexperienced raters when using SADs. The SADs developed in this study were useful for obtaining more accurate, precise and reliable assessments of loquat scab for inexperienced raters, and should be used as an aid for assessing scab in epidemiological studies or monitoring for decision-making purposes.

Keywords *Eriobotrya japonica* · Phytopathometry · Diagrammatic scale

E. González-Domínguez · J. García-Jiménez · J. Armengol (✉)
Instituto Agroforestal Mediterráneo, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain
e-mail: jarmengo@eaf.upv.es

R. B. Martins
Campus Arapiraca, Universidade Federal de Alagoas, 57309-005 Arapiraca, AL, Brazil

E. M. Del Ponte
Departamento de Fitossanidade, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, 91540-000 Porto Alegre, RS, Brazil

S. J. Michereff
Departamento de Agronomia, Universidade Federal Rural de Pernambuco, 52171-900 Recife, PE, Brazil

Introduction

Loquat (*Eriobotrya japonica* Lindl.) is a fruit tree grown in regions with a subtropical climate (Calabrese 2006) including China, Japan and the Mediterranean basin (Caballero and Fernández 2002; Lin 2007; MAGRAMA 2013). In 2006, global production of loquat fruit was estimated to be 550,000 t with a crop area >130,000 ha (Lin 2007). China is the leading producer followed by Spain. Spain is the major exporter, and >50 % of loquat production (28,812 t) is cultivated in the southeastern region of the country (Caballero and Fernández 2002; Soler et al. 2007; MAGRAMA 2013).

Loquat scab, caused by *Fusicladium eriobotryae* (Cavara) Sacc., is the major fungal disease affecting loquat in Spain (Soler et al. 2007; Sánchez-Torres et al. 2009) and other regions of the Mediterranean basin (Caballero and Fernández 2002; Sánchez-Torres et al. 2009; Gladieux et al. 2010). The fungus infects the leaves, fruits and young branches, particularly in their early stages of development. The symptoms are generally most noticeable and serious on fruit, and are first visible as circular, chlorotic spots which increase in size and become brown to olive and velvety in appearance due to production of asexual spores. Eventually, the spots become dark brown, coalesce and can cover almost the entire surface of the fruit (Rodríguez 1983; Sánchez-Torres et al. 2009) (Fig. 1). The symptoms may appear throughout the flowering and fruit developmental stages from February to June in the northern hemisphere (Sánchez-Torres et al. 2009). In years with a warm winter and a rainy spring, the incidence of the disease on fruit can be as high as 50 % (Rodríguez 1983). Scabby fruit are unacceptable for sale, resulting in significant economic losses (Sánchez-Torres et al.

Fig. 1 Fruit of loquat with typical symptoms of scab

2009). The etiology and pathogen biology of loquat scab in Spain have been studied recently (Sánchez-Torres et al. 2009; González-Domínguez et al. 2013), but the epidemiology of the disease remains poorly understood and no standardized disease assessment method is available.

While estimates of incidence can be accurate and easy to obtain, estimates of severity may be affected by the inherent ability of the rater (Campbell and Madden 1990). However it has been shown that training, experience and use of aids during assessments, such as standard area disease diagrams (SADs), lead to improved accuracy (closeness of the estimate to the true value) and reliability (the extent to which the same measurements of individuals obtained under different conditions yield similar results) of the estimates (Kranz 1988; Madden et al. 2007).

Ideally, disease assessment methods should be evaluated to ensure accuracy and reliability. Intra-rater reliability (repeatability) is the consistency of estimates by the same rater evaluated at different times, and inter-rater reliability (reproducibility) is the consistency of estimates by different raters. The percentage of leaf or fruit area affected by disease as determined by image analysis is usually considered a "true" severity (Madden et al. 2007; Bock et al. 2010). Severity can be estimated directly with or without assessment aids or using disease scales such as the one proposed by Horsfall and Barratt (Kranz 1988; Campbell and Madden 1990; Madden et al. 2007).

SADs, also known as diagrammatic scales (Godoy et al. 1997) or disease diagrams (Madden et al. 2007), are pictorial diagrams that depict the true proportion of damage (usually disease severity) on individual sampling units (quadrants, whole plants, leaves, fruit, tubers, etc.). Although it has long been stated that the use of SADs led to improved accuracy and precision of visual estimates (James 1971; Horsfall and Cowling 1978; Kranz 1988), only recently have studies provided evidence (some with statistical support) that estimates by individual raters are usually more accurate and precise when made with the aid of SADs compared to unaided estimates (Corrêa et al. 2009; Michereff et al. 2009; Lima et al. 2011; Spolti et al. 2011; Capucho et al. 2011; Bardsley and Ngugi 2013). Additionally, some studies have investigated the effect of experience of the rater on both accuracy and precision, and inexperienced raters tended to respond more to the use of SADs compared to experienced raters (Michereff et al. 2000;

Nita et al. 2003; Godoy et al. 2006; Bock et al. 2009, 2013; Pedroso et al. 2011; Sachs et al. 2011; Klosowski et al. 2013; Yadav et al. 2013).

The objectives of this study were to: 1) develop a SAD set as an assessment aid for estimating loquat severity based on a sampling of diseased fruit from the field, and 2) evaluate the effect of the SAD set and rater experience on the accuracy, precision and reliability of scab severity estimates.

Materials and methods

Loquat fruit and image acquisition

Two hundred diseased loquat fruit (cv. Algerie) were collected in May 2011 from an unsprayed loquat orchard in Callosa d'En Sarrià (Alicante province, Southeastern Spain). Cv. Algerie has previously been reported as highly susceptible to loquat scab (Sánchez-Torres et al. 2009). Fruit were mature and had a wide range of scab severity (<1 to >90 %). An experienced rater selected a subsample of 50 fruit representing the range of severity. This subset was used for all the assessments in this work. Digital images were taken from one lateral view of each fruit using a digital camera (Nikon D-5000, 12.3 megapixels) mounted 30 cm from the fruit.

Image analysis

Each image was analyzed using Assess V2.0 (Image Analysis Software for Plant Disease Quantification, American Phytopathological Society, St Paul, MN, USA (Lamari 2008)). The true severity was based on the Hue-Saturation-Intensity (HSI) colour model, which was used to determine the total area and the diseased area of the fruit in pixels. Threshold levels for healthy fruit and diseased areas were set accordingly and recorded for each fruit image.

Construction and evaluation of the SAD set

The maximum true severity determined by image analysis was 99.7 %. Because of the wide range of severity, the SAD set was constructed with eight incremental severity levels (2, 8, 16, 24, 48, 64, 86 and 98 %) with the diagrams represented in black and white (Fig. 2). Using a standard image of a fruit outline, the lesions were painted manually with shapes and spatial patterns reflecting the actual diseased fruit, which included prominent lesions around the peduncle region of the fruit.

For evaluation purposes, each image of the 50-fruit set was inserted as an individual slide in a Microsoft Powerpoint file. A group of 20 raters (10 'experienced' and 10 'inexperienced') assessed the same set of images. Raters were classified as experienced if they had received previous formal training and practice in disease severity assessment, and were familiar with the symptoms of loquat scab. Inexperienced raters had no formal training or familiarity with plant disease symptoms. For each group, 50 digital images of the diseased fruits, in a random order, were projected one at a time for 15 s. The first assessment was done without the SADs. All raters received the same information on how to identify symptoms of loquat scab and visually estimated the percentage of diseased area related to total fruit area for each image. Two weeks later, the same 20 raters estimated disease severity, but with the SADs as an aid, receiving instructions on how to use them when estimating severity to the nearest percent (i.e. use those values as reference points to help with their estimation). After a further 2-week interval, the raters made a final assessment of the fruit using the SADs (the images were displayed at random for the SADs assessments).

Data analyses

Precision, bias, accuracy, agreement and inter- and intra-rater reliability of estimates were compared without and with the use of the SADs for both inexperienced and experienced raters. Lin's concordance correlation (LCC) analysis (Lin 1989) was calculated for testing the agreement between estimate and true severity for each rater. The LCC coefficient (ρ_c) combines measures of accuracy and precision to assess the fit of pairs of observations to the line of concordance (45° = perfect concordance). The concordance correlation coefficient is calculated from the Pearson correlation coefficient (r), an indicator of precision, and the bias correction factor (C_b), which measures accuracy and is the distance to the 45° line of concordance. The bias correction factor (C_b) is calculated from μ (location bias, or height shift relative to the perfect relationship where 0 = perfect relationship between x and y) and υ (scale bias, or slope shift where 1 = perfect relation between x and y), which are derived from the means and standard deviations of x and y,

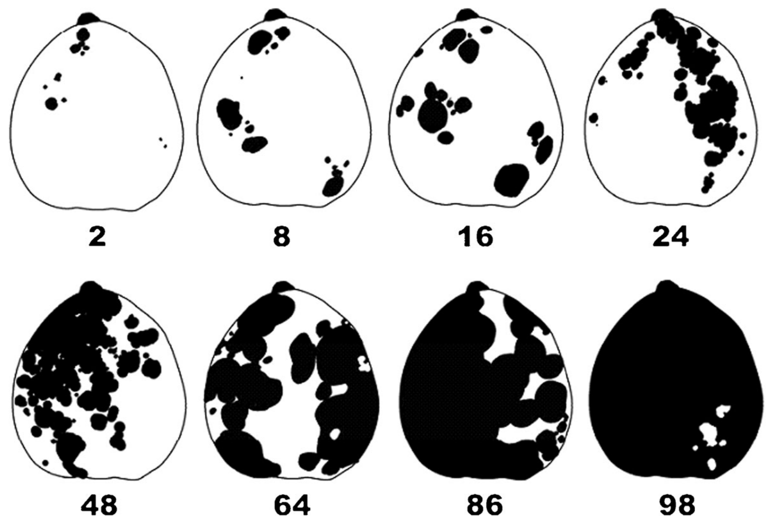

Fig. 2 The standard area diagram (SAD) set developed as an aid for assessment of scab on fruit of loquat. Each value represents the percentage of loquat scab severity on that image

respectively. Thus, the LCC coefficient provides a method to judge agreement with true values and has been previously used to judge agreement in plant disease assessments (Nita et al. 2003; Madden et al. 2007) including estimates with or without SADs (Spolti et al. 2011; Yadav et al. 2013).

Following Yadav et al. (2013), the overall effect of the use of SADs and of the experience of the rater, with or without SADs, were statistically analyzed using bootstrap analysis to calculate the mean and respective 95 % CIs on the difference between the group means for each statistics (a test of equivalence). In the analyses, 10,000 balanced bootstrap samples were taken and the 95 % CIs were calculated on the difference between the groups, so that if the CIs spanned zero, there was no significant difference (P=0.05).

The LCC statistics were used to measure intra-rater reliability for raters who repeated severity assessments with SADs. The inter-rater reliability was assessed based on the intra-class correlation coefficient (ICC) (Shrout and Fleiss 1979), which, unlike most other correlation measures, operates on data structured as groups, rather than data structured as paired observations (e.g. Pearson correlation coefficient), which are only relative measures. For our ICC analysis, the model was assumed to be two-way, with absolute agreement and single measures (Shrout and Fleiss 1979). The effect of the use of SADs on the inter-rater reliability for each assessment time was measured based on the confidence interval of the ICC estimated by the model.

All statistical analyses were calculated in R (R Core Team 2013). The epi.ccc function of the epiR package (Stevenson 2012) was used to obtain Lin's CC statistics. The built-in boot.sample R function was used for the hypothesis test. The ICC was calculated with the icc function of the irr R package (Gamer et al. 2012).

Results

Effect of SADs and rater experience on accuracy and precision

The use of the SADs resulted in improved concordance (ρ_c), accuracy (C_b) and precision (r), and reduced absolute error of the estimates compared to unaided assessments (irrespective of experience) over all raters (Fig. 3). Estimates of severity with SADs were closer to the concordance line compared to estimates without use of SADs (Fig. 3a, b). Absolute errors decreased with the use of SADs, with most values falling within ±20 %. Both without and with the use of SADs, the absolute errors of the estimates were mostly negative, suggesting a tendency to underestimate severity, particularly noticeable in the range 40 to 80 % (Fig. 3c, d).

Lin's CC statistics improved significantly with the use of SADs for the group of inexperienced raters (Table 1), with better accuracy (C_b=0.93 and 0.98 without and with SADs, respectively), precision (r=0.88 without and r=0.96 with SADs, respectively) and agreement (ρ_c without SADs = 0.82; ρc with SADs = 0.95, respectively). However, no improvements in Lin's CC statistics were found when experienced raters used the SADs (Table 1).

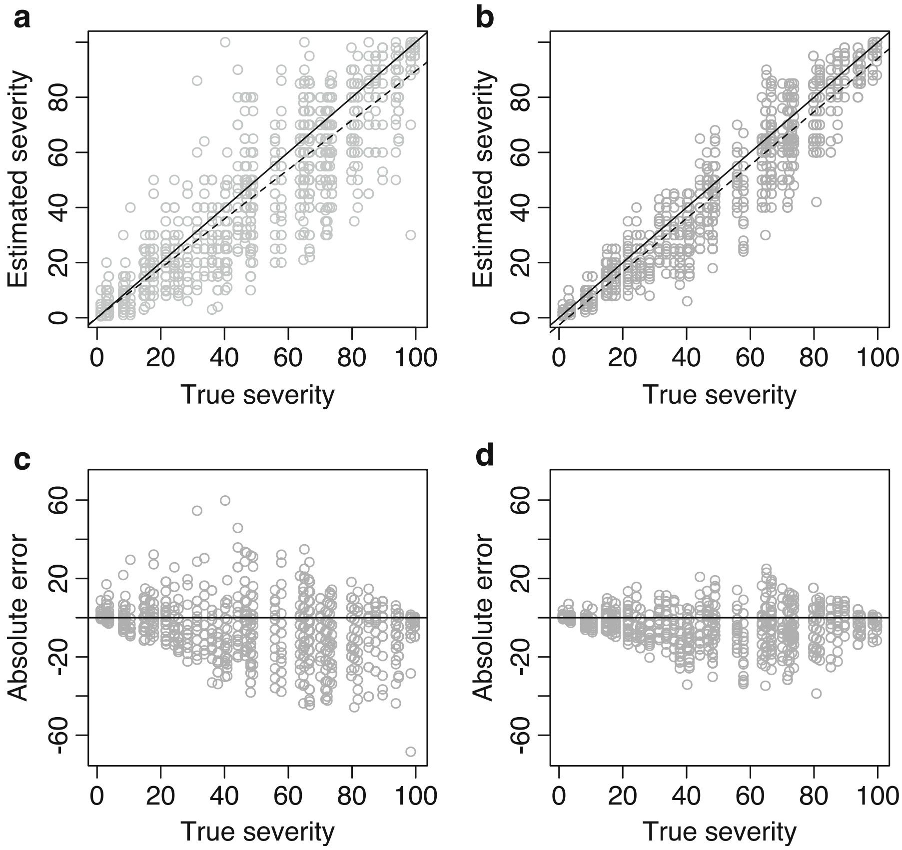

Fig. 3 Relationship between the estimates and true severity (**a**, **b**) and the absolute error (estimate minus true severity) (**c**, **d**) of assessments of a set of 50 images of scab-diseased fruit of loquat by 20 raters without (**a**, **c**) and with (**b**, **d**) the use of a standard area diagrams (SADs). *Dashed line* (**a**, **b**) represents the regression line

The greatest improvement in accuracy of estimates with SADs compared to without SADs was noted for the group of inexperienced raters (Fig. 4). The biggest gains in agreement, bias and precision were obtained for raters with the poorest estimates without the use of SADs (Fig. 4a–c). Intriguingly, some experienced raters increased bias (gain was negative) thus resulting in reduced agreement when using the SADs compared to estimates initially made without the SADs (Fig. 4).

Rater experience had a significant effect on precision (r) but not the accuracy (C_b) for estimates made without SADs (Table 2). With the use of SADs, inexperienced raters were significantly more accurate and precise compared to experienced raters, especially during the second assessment using SADs. Accuracy (C_b), precision (r) and agreement (p_c) differed significantly between the experienced and inexperienced raters based on the bootstrap analysis, with values, in the case of inexperienced raters in the range 0.97 to 0.99.

Inter-rater and intra-rater reliability

Inter-rater reliability, based on the intra-class correlation coefficient, was improved for inexperienced raters with use of SADs (ρ=0.71 without SADs and ρ=0.93 and ρ=0.96 for the first and second assessment using SADs, respectively) (Table 3). Experienced raters showed good inter-rater reliability (ρ>0.94) irrespective of the use of SADs.

Good intra-rater reliability was obtained by both experienced and inexperienced raters when estimates were made using the SADs (ρ_c>0.92). There was no effect of experience on intra-rater reliability (mean ρ_c= 0.95 and 0.95 for inexperienced and experienced raters, respectively) (Table 4).

Discussion

The SAD set of eight black and white loquat scab diseased fruit images covered the range 2 % to 98 %

Table 1 Effect of use of SADs as an assessment aid on the bias, precision and overall agreement of estimates of scab severity on loquat fruit made by raters with or without experience in disease severity assessment

Experience	LCC statistic	Means		95 % CI[a] of the difference between means
		No SAD aid	With SAD aid	
Inexperienced	Scale-shift (υ)[b]	1.01	1.00	−0.059, 0.091
	Location-shift (μ)[c]	0.18	0.09	−0.120, 0.295
	Bias correction factor (C_b)[d]	0.93	0.98	**−0.094, −0.014**
	Correlation coefficient (r)[e]	0.88	0.96	**−0.117, −0.052**
	Concordance coefficient (ρc)[f]	0.82	0.95	**−0.175, −0.089**
Experienced	Scale-shift (υ)	1.04	0.99	−0.007, 0.106
	Location-shift (μ)	0.19	0.20	−0.133, 0.120
	Bias correction factor (C_b)	0.96	0.97	−0.032, 0.025
	Correlation coefficient (r)	0.95	0.95	−0.015, 0.008
	Concordance coefficient (ρ_c)	0.92	0.93	−0.042, 0.028

[a] Bootstrap calculated difference between means and confidence intervals (CIs). If the CIs embrace zero, difference is not significant at the 5 % level. Bold numbers represent significance of the difference

[b] Scale or slope shift relative to the perfect relationship (1 = perfect relation between x and y)

[c] Location or height shift relative to the perfect relationship (0 = perfect relation between x and y)

[d] Bias correction factor that measures how far the best-fit line deviates from a line at 45°. No deviation from the 45° line occurs when C_b=1. C_b is calculated from υ and μ and is a measure of accuracy

[e] Correlation coefficient (r) that measures precision

[f] Lin's concordance correlation coefficient (ρ_c) combines both precision (r) and accuracy (C_b) (ρ_c=r.C_b) to measure agreement with the true value (Lin 1989)

severity and exhibited typical symptom patterns of loquat scab (small spots initially, that expand and coalesce covering almost the entire fruit surface; Sánchez-Torres et al. (2009)). This severity range is commonly observed in cv. Algerie, the most widely grown cultivar in Spain and known to be highly susceptible to *F. eriobotryae* (Sánchez-Torres et al. 2009). More realistically, the diagrams, although depicting only the side view of the fruit, display the typical prominent lesions at the peduncle region of more severely affected fruit, a pattern commonly associated with splash-dispersed pathogens (Bock et al. 2011).

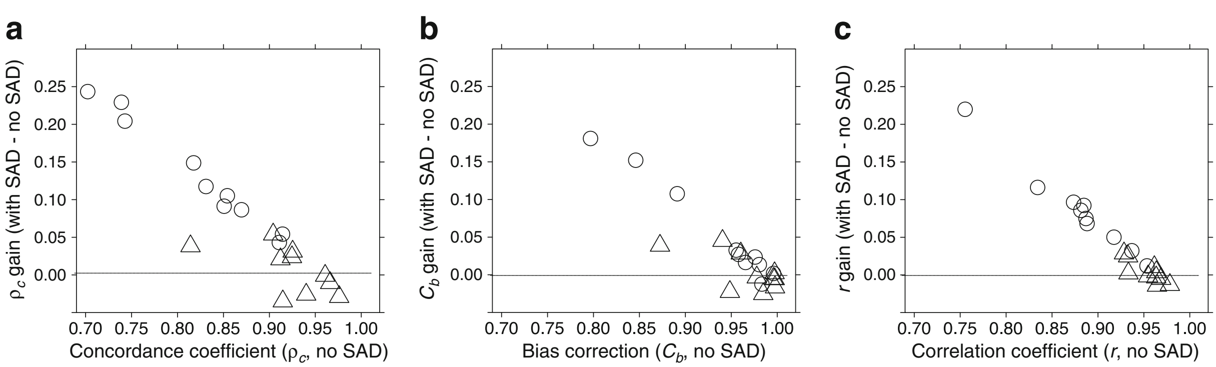

Fig. 4 The relationship between gain (difference between the estimate with and without use of a standard area diagram, SAD) for measures of accuracy, precision and agreement of visual severity estimates made by 10 inexperienced (*circles*) and 10 experienced (*triangles*) raters for a set of 50 images of scab-diseased fruit of loquat (**a**, agreement, measured by Lin's concordance correlation coefficient; **b** accuracy measured by the bias correction factor; and **c** precision, measured by the correlation coefficient)

Table 2 Effect of rater experience on the bias, precision and overall agreement of estimates of scab severity on loquat fruit made by ten raters either unaided or aided by a standard area diagram (SAD)

Assessment	LCC variable	Means		95 % CI[a] of the difference between means
		Not experienced	Experienced	
No SAD	Scale-shift (υ)[b]	1.01	1.04	−0.103, 0.053
	Location-shift (μ)[c]	0.18	0.19	−0.240, 0.211
	Bias correction factor (C_b)[d]	0.93	0.96	−0.078, 0.013
	Correlation coefficient (r)[e]	0.88	0.95	**−0.108, −0.039**
	Concordance coefficient (ρ_c)[f]	0.82	0.92	**−0.151, −0.049**
SAD first assessment	Scale-shift (υ)	1.00	0.99	−0.050, 0.066
	Location-shift (μ)	0.09	0.20	**−0.200, −0.007**
	Bias correction factor (C_b)	0.98	0.97	−0.001, 0.037
	Correlation coefficient (r)	0.96	0.95	−0.001, 0.015
	Concordance coefficient (ρ_c)	0.95	0.93	**0.003, 0.048**
SAD second assessment	Scale-shift (υ)	1.02	0.98	**0.011, 0.071**
	Location-shift (μ)	0.07	0.20	**−0.212, −0.035**
	Bias correction factor (C_b)	0.99	0.97	**0.002, 0.043**
	Correlation coefficient (r)	0.98	0.95	**0.012, 0.031**
	Concordance coefficient (ρ_c)	0.97	0.93	**0.018, 0.068**

[a] Bootstrap calculated difference between means and confidence intervals (CIs). If the CIs embrace zero, difference is not significant at the 5 % level. Bold numbers represent significance of the difference

[b] Scale or slope shift relative to the perfect relationship (1 = perfect relation between x and y)

[c] Location or height shift relative to the perfect relationship (0 = perfect relation between x and y)

[d] Bias correction factor that measures how far the best-fit line deviates from a line at 45°. No deviation from the 45° line occurs when C_b=1. C_b is calculated from υ and μ and is a measure of accuracy

[e] Correlation coefficient (r) that measures precision

[f] Lin's concordance correlation coefficient (ρ_c) combines both precision (r) and accuracy (C_b) ($\rho_c = r.C_b$) to measure agreement with the true value (Lin 1989)

The number of diagrams (eight) used to depict the severity range (2–98 %) was similar to the number used in SADs of other pathosystems, which is of practical use (Corrêa et al. 2009, Yadav et al. 2013). The severity increment was nonlinear because more reference values (five) were placed in a range where loquat scab severity levels are typically more common (<50 %) (results of this study; Sánchez-Torres et al. 2009).

Although the SADs improved the accuracy and reliability of the estimates by inexperienced raters, there was an overall tendency to underestimate severity in the 40–80 % range of severity. Most reports in the literature show that inexperienced raters tended to overestimate severity, especially at lower severities (<20 %) (Forbes and Korva 1994; Diaz et al. 2001; Leite and Amorim 2002; Spósito et al. 2004; Bock et al. 2008, 2009). However, this is not the first study to report a slight tendency of raters to underestimate severity at mid to high disease severity (Michereff et al. 1998, 2000;

Table 3 Inter-rater reliability (reproducibility) of visual estimates of scab severity on 50 loquat fruit by 20 raters either unaided (one-time assessment) or aided by SADs measured by the intra-class correlation coefficient (ICC). Two assessments were made using the SADs with a 2-week interval

Assessment	Intra-class correlation coefficient, ρ (95 % CI)	
	Inexperienced raters (n=10)	Experienced raters (n=10)
No SAD	0.71 (0.59–0.80)	0.94 (0.91–0.96)
SAD aid 1st assessment	0.93 (0.91–0.95)	0.94 (0.92–0.96)
SAD aid 2nd assessment	0.96 (0.94–0.97)	0.95 (0.93–0.97)

Table 4 Intra-rater reliability (repeatability), measured by Lin's concordance correlation coefficient (ρ_c), for estimates of severity on 50 diseased loquat fruit by two groups of ten raters either with or without experience but using SADs during two consecutive SAD-aided assessments

Rater	Experience	
	No (n=10)	Yes (n=10)
1	0.97	0.97
2	0.95	0.95
3	0.97	0.97
4	0.97	0.97
5	0.94	0.94
6	0.97	0.97
7	0.95	0.94
8	0.94	0.94
9	0.92	0.92
10	0.97	0.97
Mean	0.955	0.954
Diff. between means[a]	0.001 (0.0077)	
95%CI of the difference[b]	−0.0142, 0.0162	

[a] Bootstrap calculated difference between means

[b] If the CIs embrace zero, difference is not significant at P=0.05

Gomes et al. 2004; Spolti et al. 2011). This underestimation could be because of the disease pattern on the fruit. Spolti et al. (2011) discussed underestimation of severity of sooty blotch and flyspeck disease as possibly related to the clustering of small lesions that could not be discerned easily by raters. Moreover, Sherwood et al. (1983) also demonstrated that illusions in assessment due to lesion size and number resulted in error. If two leaves had similar severity but one had many small lesions as compared to a few larger lesions, the visual estimate for the leaf with the greater number of lesions exceeded that for the leaf with the fewer (but larger) lesions. Although this situation related to overestimation, equivalent illusions may be a cause of underestimates we observed with loquat scab. In the case of loquat scab, clustering of lesions on the fruit surface is typical of the disease (Sánchez-Torres et al. 2009), reducing the number of diseased areas at higher values of severity. The clustering of disease might affect how the symptom severity is perceived. Furthermore, the fact that the SADs have only three severity values > 50 % might also have contributed to greater negative bias in that range.

The effect of rater experience on the accuracy and precision of estimates, without or with SADs, has been studied previously (Godoy et al. 2006; Pedroso et al. 2011; Yadav et al. 2013). In an evaluation of SADs for pecan scab caused by *Fusicladium effusum* (Yadav et al. 2013), experienced raters who showed low bias and high accuracy and precision of estimates without SADs, did not respond as much as the inexperienced raters to the use of SADs. The results of the current study confirmed that experienced raters did not benefit significantly from the use of SADs.

Significant improvements in reliability of the estimates, both within and among raters, were noticed for estimates made with the SADs, in agreement with previous reports (Godoy et al. 2006; Yadav et al. 2013). The SADs developed in our study proved to be useful for obtaining more accurate, precise and reliable estimates of loquat scab for inexperienced, rather than experienced raters, and will be a valuable tool for raters as an aid in estimating scab severity in epidemiological studies or for monitoring disease for decision-making purposes.

Acknowledgments This work was funded by Cooperativa Agrícola de Callosa d'En Sarrià (Alicante, Spain). Financial support was provided by CAPES (Brazil) for S. J. Michereff at the Universidad Politécnica de Valencia (Spain) and is gratefully acknowledged. The authors thank colleagues and students who assessed the diseased fruits images.

References

Bardsley, S. J., & Ngugi, H. K. (2013). Reliability and accuracy of visual methods to quantify severity of foliar bacterial spot symptoms on peach and nectarine. *Plant Pathology, 62*, 460–474.

Bock, C. H., Parker, P. E., Cook, A. Z., & Gottwald, T. R. (2008). Characteristics of the perception of different severity measures of citrus canker and the relations between the various symptom types. *Plant Disease, 92*, 927–939.

Bock, C. H., Parker, P. E., Cook, A. Z., Riley, T., & Gottwald, T. R. (2009). Comparison of assessment of citrus canker foliar symptoms by experienced and inexperienced raters. *Plant Disease, 93*, 412–424.

Bock, C. H., Poole, G., Parker, P. E., & Gottwald, T. R. (2010). Plant disease severity estimated visually by digital photography and image analysis and by hyperspectral imaging. *Critical Reviews in Plant Sciences, 29*, 59–107.

Bock, C. H., Gottwald, T. R., & Parker, P. E. (2011). Distribution of canker lesions on the surface of diseased grapefruit. *Plant Pathology, 60*, 986–991.

Bock, C. H., Wood, B. W., van den Bosch, F., Parnell, S., & Gottwald, T. R. (2013). The effect of Horsfall-Barratt category size on the accuracy and reliability of estimates of pecan scab severity. *Plant Disease, 97*, 797–806.

Caballero, P., & Fernández, M. A. (2002). Loquat, production and market. *Options Méditerranéennes, 58*, 11–20.

Calabrese, F. (2006). Origen de la especie. In M. Agustí, C. Reig, & P. Undurraga (Eds.), *El Cultivo del Níspero Japonés* (pp. 17–26). Valencia: Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile and Universidad Politécnica de Valencia.

Campbell, C. L., & Madden, L. V. (1990). *Introduction to plant disease epidemiology*. New York: Wiley. 532 pp.

Capucho, A. S., Zambolim, L., Duarte, H. S. S., & Vaz, G. R. O. (2011). Development and validation of a standard area diagram set to estimate severity of leaf rust in *Coffea arabica* and *C. canephora*. *Plant Pathology, 60*, 1144–1150.

Corrêa, F. M., Bueno Filho, J. S. S., & Carmo, M. G. F. (2009). Comparison of three diagrammatic keys for the quantification of late blight in tomato leaves. *Plant Pathology, 58*, 1128–1133.

Diaz, C. G., Bassanezi, R. B., & Bergamin Filho, A. (2001). Desenvolvimento e validação de uma escala diagramática para *Xanthomonas axonopodis* pv. *phaseoli* em feijoeiro. *Summa Phytopathologica, 27*, 35–39.

Forbes, G. A., & Korva, J. T. (1994). The effect of using Horsfall-Barratt scale on precision and accuracy of visual estimation of potato late blight severity in the field. *Plant Pathology, 43*, 675–682.

Gamer, M., Lemon, J., & Fellows, I. (2012). *irr: Various Coefficients of Interrater Reliability and Agreement*. R Package Version 0.84.

Gladieux, P., Caffier, V., Devaux, M., & Le Cam, B. (2010). Host-specific differentiation among populations of *Venturia inaequalis* causing scab on apple, pyracantha and loquat. *Fungal Genetics and Biology, 47*, 511–521.

Godoy, C. V., Carneiro, S. M. T. P. G., Iamauti, M. T., Dalla Pria, M., Amorim, L., Berger, R. D., & Bergamim Filho, A. (1997). Diagramatic scale for bean: development and validation. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz, 104*, 336–345.

Godoy, C. V., Koga, L. J., & Canteri, M. G. (2006). Diagrammatic scale for assessment of soybean rust severity. *Fitopatologia Brasileira, 31*, 63–68.

Gomes, A. M. A., Michereff, S. J., & Mariano, R. L. R. (2004). Development and validation of a diagramatic key for Cercospora leaf spot of lettuce. *Summa Phytopathologica, 30*, 38–42.

González-Domínguez, E., Rossi, V., Armengol, J. & García-Jiménez, J. (2013). Effect of environmental factors on mycelial growth and conidial germination of *Fusicladium eriobotryae*, and the infection of loquat leaves. *Plant Disease, 97*, 1331–1338.

Horsfall, J. G., & Cowling, E. B. (1978). Pathometry: the measurement of plant disease. In J. G. Horsfall & E. B. Cowling (Eds.), *Plant disease: An advanced treatise* (Vol. II, pp. 20–136). New York: Academic.

James, W. C. (1971). An illustrated series of assessment keys for plant diseases, their preparation and usage. *Canadian Plant Disease Survey, 51*, 39–65.

Klosowski, A. C., Ruaro, L., Bespalhok Filho, J. C., & May De Mio, L. L. (2013). Proposal and validation of diagrammatic scale for assessment of orange rust of sugarcane. *Tropical Plant Pathology, 38*, 166–171 (in Portuguese).

Kranz, J. (1988). Measuring plant disease. In J. Kranz & J. Rotem (Eds.), *Experimental techniques in plant disease epidemiology* (pp. 35–50). New York: Springer.

Lamari, L. (2008). *ASSESS 2.0: Image analysis software for plant disease quantification*. St. Paul: APS Press.

Leite, R. M. V. B. C., & Amorim, L. (2002). Elaboração e validação de escala diagramática para mancha de Alternaria em girassol. *Summa Phytopathologica, 28*, 14–19.

Lima, G. S. A., Assunção, I. P., Martins, R. B., Santos, H. V., & Michereff, S. J. (2011). Development and validation of a standard area diagram set for assessment of Alternaria spot on the cladodes of the prickly pear cactus. *Journal of Plant Pathology, 93*, 691–695.

Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics, 45*, 255–268.

Lin, S. Q. (2007). World loquat production and research with special reference to China. *Acta Horticulturae, 750*, 37–44.

Madden, L. V., Hughes, G., & van den Bosch, F. (2007). *The study of plant disease epidemics*. St. Paul: APS Press. 421 pp.

MAGRAMA. (2013). *Anuario de Estadística 2012*. Madrid: Ministerio de Medio Agricultura, Alimentación y Medio Ambiente. 891 pp.

Michereff, S. J., Pedrosa, R. A., Noronha, M. A., Martins, R. B., & Silva, F. V. (1998). Escala diagramática e tamanho de amostras para avaliação da severidade da mancha parda da mandioca (*Cercosporidium henningsii*). *Agrotrópica, 3*, 143–148.

Michereff, S. J., Maffia, L. A., & Noronha, M. A. (2000). Diagrammatic key for assessment of yam leaf blight severity. *Fitopatologia Brasileira, 25*, 612–619 (in Portuguese).

Michereff, S. J., Noronha, M. A., Lima, G. S. A., Albert, I. C. L., Melo, E. A., & Gusmão, L. O. (2009). Diagrammatic scale to assess downy mildew severity in melon. *Horticultura Brasileira, 27*, 76–79.

Nita, M., Ellis, M. A., & Madden, L. V. (2003). Reliability and accuracy of visual estimation of Phomopsis leaf blight of strawberry. *Phytopathology, 93*, 995–1005.

Pedroso, C., Lage, D. A. C., Henz, G. P., & Café-Filho, A. C. (2011). Development and validation of a diagrammatic scale for estimation of anthracnose on sweet pepper fruits for epidemiological studies. *Journal of Plant Pathology, 93*, 219–225.

R Core Team. (2013). *R: A language & environment for statistical computing*. Vienna: R Foundation for Statistical Computing. http://www.R-project.org/.

Rodríguez, A. (1983). *El Cultivo del Níspero en el Valle de Algar-Guadalest*. Alicante: Sociedad Cooperativa de Crédito de Callosa d'En Sarrià, Callosa d'En Sarrià. 262 pp.

Sachs, P. J. D., Neves, C. C. S. V. J., Canteri, M. G., & Sachs, L. G. (2011). Diagrammatic scale for assesment of the Phaeosphaeria leaf spot severity in maize. *Summa Phytopathologica, 37*, 202–204.

Sánchez-Torres, P., Hinarejos, R., & Tuset, J. J. (2009). Characterization and pathogenicity of *Fusicladium eriobotryae*, the fungal pathogen responsible for loquat scab. *Plant Disease, 93*, 1151–1157.

Sherwood, R. T., Berg, C. C., Hoover, M. R., & Zeiders, K. E. (1983). Illusions in visual assessment of *Stagonospora* leaf spot of orchardgrass. *Phytopathology, 73*, 173–177.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 86*, 420–428.

Soler, E., Martínez-Calvo, J., Llácer, G., & Badenes, M. L. (2007). Loquat in Spain: production & market. *Acta Horticulturae, 750*, 45–47.

Spolti, P., Schneider, L., Sanhueza, R. M. V., Batzer, J. C., Gleason, M. L., & Del Ponte, E. M. (2011). Improving sooty blotch and flyspeck severity estimation on apple fruit with the aid of standard area diagrams. *European Journal of Plant Pathology, 129*, 21–29.

Spósito, M. B., Amorim, L., Belasque, J., Bassanezi, R. B., & Aquino, R. (2004). Elaboração e validação de escala diagramática para avaliação da severidade da mancha preta em frutos cítricos. *Fitopatologia Brasileira, 29*, 81–85.

Stevenson, M. (2012). *Epir: An R Package for the Analysis of Epidemiological Data*. R package version 0.9-43.

Yadav, N. V. S., de Vos, S. M., Bock, C. H., & Wood, B. W. (2013). Development and validation of standard area diagrams to aid assessment of pecan scab symptoms on fruit. *Plant Pathology, 62*, 325–335.