Lesion distribution, multiple infection, and the logistic increase of plant disease

(negative binomial distribution/contagious distribution)

PAUL E. WAGGONER AND SAUL RICH

Department of Plant Pathology and Botany, The Connecticut Agricultural Experiment Station, New Haven, Connecticut 06504

Contributed by Paul E. Waggoner, March 16, 1981

ABSTRACT The lesions of few of the 112 observed cases of plant disease were randomly distributed; most were fitted by the negative binomial function with exponent ≈ 1 . The proportion of healthy hosts generally equaled the frequency of zero in the negative binomial. The same course of an epidemic is calculated for the logistic increase of infected plants or for plants infected by propagules produced in proportion to lesions and distributed with exponent equal 1.

An assumption of the random distribution of propagules is the foundation of the two most important mathematical models in plant pathology: the multiple infection transformation, and the logistic law. We examine this foundation.

The multiple infection transformation, which was founded upon W. R. Thompson's relation between the numbers of parasites and of hosts parasitized, was introduced into plant pathology by Gregory (1). It predicts that the increase in the number N of diseased hosts per addition of effective propagules n is proportional to the fraction $(K - N)/K$ of the population K that is healthy at the time of dissemination:

$$dN/dn = (K - N)/K. \quad [1]$$

The essential assumptions are that all n propagules infect, that they are randomly distributed over the K hosts, and that those propagules that cause multiple infections do not increase N . The familiar multiple infection transformation is derived from Eq. 1:

$$(1 - x) = \exp(-n/K) \quad \text{or} \quad n = -K \ln(1 - x) \quad [2]$$

which relates the n propagules to the infected portion N/K or x of K hosts.

The second model, the logistic law, first written by Verhulst in 1845 and demonstrated in plant pathology by van der Plank (2), predicts the increase dN/dt in diseased hosts with time t . The production of propagules is assumed to be proportional to the number of diseased hosts and, again, their rate of infecting healthy hosts is made $(K - N)/K$ on the assumption that they are randomly distributed.

$$dN/dt = rN(K - N)/K \quad [3]$$

where r is the proportionality factor. Eq. 3 produces the familiar

$$\ln[x/(1 - x)] - \ln[x_0/(1 - x_0)] = r(t - t_0) \quad [4]$$

which leads us to estimate the rate r of increase of infection in a field by fitting a straight line to logits, $\ln[x/(1 - x)]$ (curve 3 in Fig. 1).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

FIG. 1. The increase of diseased hosts N on logit scale. Curves are identified by equation number followed by value for k in parentheses if appropriate: the logistic law (Eq. 3); Eq. 3 modified for propagule formation proportional to lesions (Eq. 9); increase when hosts infected per propagule is modified for contagion (Eq. 8); as Eq. 8 and propagule formation proportional to lesions (Eq. 10). The r is 1.

We examined the assumption of random distribution of effective propagules by counting lesions per leaflet, leaf, or plant. Because we dealt with discrete lesions that occupied little of the total area of the host, we equated n effective propagules with n lesions.

If lesions are randomly distributed as assumed, their frequency distribution will be fitted by the Poisson function. The $\exp(-n/K)$ in Eq. 2 is the relative frequency of zeroes in a Poisson with mean m equal to n/K . Thus, we tested the assumption underlying the models by comparing the Poisson to the frequency distribution of the lesions of a range of diseases.

At the outset, however, one expects that propagules of a contagious disease will not be randomly distributed. Instead, the expected frequency distribution will be "contagious" or "over-dispersed" with variance larger than the mean and fitted by the negative binomial distribution. Our procedure is describing the negative binomial, showing that it rather than the Poisson fits a range of diseases, and discussing required changes in the logistic law and multiple infection transformation.

The probability P_y of observing any positive integer y depends upon the mean m and the positive exponent k :

$$P_y = \frac{(k + y - 1)!}{y! (k - 1)!} \left(\frac{1}{q}\right)^k \left(\frac{p}{q}\right)^y \quad [5]$$

where p is m/k and q is $p + 1$. When the distribution is over-dispersed, k is small. As over-dispersion decreases, k approaches infinity, and the distribution converges to the Poisson (3, 4).

Because disparate mechanisms produce data conforming to the negative binomial (3, 4), we cannot use agreement between

observations and the function to justify a particular model. On the other hand, we know that some mechanisms of infection conform to models producing the negative binomial, we can predict that those infections will be distributed as the negative binomial, and we can test our prediction by counting lesions.

Because a mixture of Poisson distributions conforms to the negative binomial if the means of the Poisson distributions vary appropriately, lesions produced by a series of spore showers may be distributed as the negative binomial. Rouse *et al.* (5) hypothesized that varying leaf sizes, spore deposition, and spore success produced the negative binomial distribution of colonies of *Erysiphe graminis* f. sp. *tritici* on wheat.

If the expected frequency of success of propagules varies within an observed distribution, the result may be a negative binomial. Gregory (1) attributed the nonrandom distribution of lesions caused by *Gymnosporangium* to varying susceptibility, and varying microclimate could have a similar consequence.

If the presence of one individual in an area increases the chance of others falling into that area, a negative binomial will fit best (6). When infection is spread by splashing drops of rain with some carrying many propagules from a lesion and some carrying none, the chance of a second infection is higher if a first is present, and a negative binomial distribution may be produced. Campbell and Pennypacker (7) suggested that large pieces of decaying tissue caused clusters of lesions of *Rhizoctonia solani* on beans and a negative binomial distribution.

Finally, birth, death, and immigration can cause a negative binomial. The negative binomial distribution of cankers per tree was attributed to primary infection of an orchard by *Nectria galligena*, its multiplication within the orchard, and pruning of cankers (8). Reproduction within a small region was credited for the negative binomial distribution of lesions of *Xanthomonas compestris* on cabbage (9).

MATERIALS AND METHODS

Mixture of Distributions. Although influences such as variations in resistance or environment within a set of observations may have acted, we expect that departures from random distribution of the first six pathogens of Table 1 were largely caused by the mixture of spore showers. The dates of observations and the numbers of hosts per frequency distribution are in parentheses in the following.

Ash rust is caused on *Fraxinus americana* by urediospores of *Puccinia peridermiospora* blown from marsh grass *Spartina* spp.; secondary spread does not occur on ash. The lesions were counted on four or five outer leaflets of leaves on trees 75–200 m from a marsh covered by *Spartina* in Guilford, CT (14 and 28 June 1980; 24–235 leaflets). Similar observations were made 10–30 km north of the marshes along Long Island Sound (11–29 June 1980; 32–235 leaflets).

Apple scab is initiated annually on *Malus pumila* by ascospores of *Venturia inaequalis* from old leaves on the ground. In Hamden, CT, the first lesions on McIntosh apples were counted on the five oldest leaves on 10 spurs on four trees (20 May 1980; 40 leaves). Again in Hamden, primary lesions were observed on spurs of McIntosh with at least one lesion (23 May 1980; 311 leaves). In New Haven, CT, lesions that were both primary and secondary were counted on a crab apple; 50 leaves came from a portion of the tree that was periodically sprinkled as the lawn was watered, and 100 came from a portion that was not (5 June 1980; 150 leaves).

Rust of corn (*Zea mays*) is initiated by urediospores of *Puccinia sorghi* from *Oxalis* spp. Urediospores likely do not overwinter in Connecticut but instead are produced in more southern fields infected earlier. In Hamden, the first pustules were counted on the top five leaves of plants of 18 varieties (15 August

1978; 10 plants). Eight days later, the more numerous pustules were counted (23 August 1978; 15 plants) (10).

Rust of shadbush is caused on *Amelanchier arborea* by teliospores of *Gymnosporangium clavipes* carried from *Juniperus virginiana*, and no secondary spread occurs on *Amelanchier*. The lesions were counted on leaves on two branches in Guilford (7 June 1980; 92 and 113 leaves). Eight days later, pustules were counted again (15 June 1980; 118 leaves).

Cedar-apple rust is caused on *Malus pumila* by teliospores of *Gymnosporangium juniperi-virginianae* carried from *Juniperus virginiana*, and no secondary spread occurs on *Malus*. The lesions were counted on leaves on eight branches of Smoothie apples in Hamden (15 and 18 August 1980; 44–88 leaves).

Frogeye leaf spot of apple is caused by conidia of *Physalospora malorum* carried from mummified fruit of *Malus pumila* to leaves, but spores are rarely produced on leaves. In Guilford, distinct lesions were counted on mature leaves of a crab apple; indistinct and new lesions were counted on mature leaves and on newly expanded leaves (31 May and 8 June 1980; 36–116 leaves). Lesions were counted on leaves of one Delicious and six McIntosh trees (26 and 31 May and 1 June 1980; 22–119 leaves).

Variation in Success. Differences in environment or variety were created by combining some preceding observations.

Presence of One Changes Chance of a Second. We expect that departures from random distribution of the early infections of the two bacterial diseases were caused largely by a drop splashed from a healthy leaf with no propagules or from a diseased leaf with many.

Angular leaf spot of cucumber (*Cucumis sativus*) is caused by drops laden with *Pseudomonas lachrymans*. On 30 July 1980, *Ps. lachrymans* was sprayed on a leaf in two regions of a field in Hamden. Nine days later, lesions appeared. Sixteen, 21, and 37 days after we inoculated the two leaves, lesions were counted on all leaves within 1 m² around each inoculated leaf and in an adjacent 1 m² (15 and 20 August and 5 September 1980; 113–132 leaves).

Bacterial spot of pepper (*Capsicum frutescens*) is caused by drops carrying *Xanthomonas vesicatoria*. Lesions were counted in Simsbury, CT (7 August 1980; 112 leaves). On 20 August, 15–20 of these leaves were scattered on four plants in four quadrants immediately irrigated for 1 hr, and then for 4 hr on the next day. Twenty-seven days later, the lesions were counted on five inoculated plants and nine adjacent plants, which were 46 cm distant (16 September 1980; 64–148 leaves).

Birth, Death, and Migration. Two diseases were chosen to illustrate departure from a random distribution caused by multiplication of a pathogen on a host and the loss of diseased organs while more propagules migrated from afar.

Black knot is caused by ascospores of *Dibotryon morbosum* washing onto new twigs of *Prunus* species. Knots were counted on *P. serotina* 77–150 cm tall in Farmington, CT (26 June 1980; 50 trees).

Blue mold of tobacco is caused by spores of *Peronospora tabacina* air-borne onto *Nicotiana tabacum* and from place to place on the plants. Lesions were counted on plants in three fields of Broadleaf tobacco in South Windsor, CT (12–13 July 1955; 261–555 plants) (11). Plants with more than 100 lesions in one field were omitted, and counts concentrated at 50, 60, . . . , were distributed through adjacent classes.

Size. Although the observed populations had no gross heterogeneities, their size inevitably varied. The effect of size was investigated in ash rust, frogeye leafspot, bacterial spot of pepper, and plum knot. The number *n* of lesions was adjusted for size by multiplying by the ratio of the average size to the size of the individual. Size was the length times width of leaflets or

leaves or the height of *Prunus*. An adjusted n was taken as an integer.

Statistical Analysis. Departure of distributions from the Poisson was tested by χ^2 equal to the degrees of freedom times the ratio of the mean square to the mean. Departure from the negative binomial was tested by a difference between observed and expected moments, the statistic U or T (4).

RESULTS

Size. Although adjustment of the lesions of 10 distributions of four diseases for the size of leaflet, leaf, or plant decreased the mean square by 2–46%, adjustment did not cause the Poisson to fit, the binomial to fail to fit, or significantly change the estimate of k . The statistics hereafter were estimated from the lesions counted.

Mixture of Distributions. Sixty-nine of the 82 distributions of the four rusts, of apple scab, and of frog-eye leaf spot (lines 1 through 6 in Table 1) were not fitted by the Poisson. That is, 69 χ^2 were significant at the 5% level, indicating over-dispersion or contagion. The 13 fitted by the Poisson had means <1 . Sixty-seven of the 69 distributions not fitted by the Poisson were fitted by the negative binomial with k from 0.01–4.0.

Variation in Success. The effect of variation in resistance is demonstrated with corn rust. The lesions, on 18 August 1978, of 5 of the 18 varieties were fitted by the Poisson. The mean lesions on five leaves of these varieties was <2 ; on a variety susceptible at infection, it was 24. Nevertheless, when the five random distributions were pooled, the distribution differed from the Poisson and was fitted by the negative binomial with $k = 1.8$.

The effect of variation in environment is demonstrated with apple scab on crabapple. On 5 June 1980, lesions on 50 leaves wetted by sprinkling had a mean of 1.9, and their distribution was fitted by the negative binomial with $k = 1.2$. The mean on 100 leaves on branches that were not sprinkled was only 0.1. For the pooled distribution, k was only 0.2.

Presence of One Changes Chance of a Second. Twenty-five of the 26 distributions of the lesions of the splash-borne bacteria (lines 7 and 8, Table 1) were not fitted by the Poisson function although several had means <1 . Twenty-four of the 25 distributions, however, were fitted by the negative binomial with k 0.01–0.33. Only 7 days after the primary lesions appeared on the cucumbers, the frequency distribution required k only 0.01–0.07.

Birth, Death, and Migration. Although no distribution of black knot or blue mold lesions per plant (lines 9 and 10 in Table 1) was fitted by the Poisson, two were fitted with $k = 0.4$ and 2.4.

DISCUSSION

We predicted that certain characteristics of each disease would cause over-dispersion of the frequency distributions—i.e., more hosts healthy and more with many lesions than random distribution of propagules and consequent lesions would produce. In fact, 98 of 112 distributions departed significantly from the Poisson. The degree of the over-dispersion is exemplified by a sample of 235 ash leaflets in which 52% rather than the expected 35% were healthy and 5% rather than 0.4% had more than four lesions. Only 5 of the 112 distributions were not fitted by the negative binomial.

Although we argued that a single mechanism largely caused a deviation from random distribution, we cannot, of course, exclude other mechanisms. For example, although deviation from random distribution of scab and frog-eye leaf spot of apple was attributed to a mixture of spore distributions, the fall of severely diseased leaves also modified the frequency distribution. Nevertheless, a mixture of frequency distributions undoubtedly struck these leaves and caused a nonrandom frequency distribution. Thus, use of diseases to exemplify mechanisms, not to claim a single mechanism, seems justified.

With the general inapplicability of the Poisson demonstrated by the range of diseases tabulated in Table 1 and elsewhere (5, 7–9), we turn to the changes required in the multiple infection transformation and the logistic law, which assume the Poisson function fits.

The multiple infection transformation, Eq. 2, relates the n propagules or subsequent lesions on K hosts to the proportion x of hosts diseased. Eq. 2 suggests a graphical test: the observations should fall on the straight diagonal line in Fig. 2 where $-\ln(1-x)$ equals the mean n/K or m . If, however, the negative binomial fits the frequency distribution, the relation between the proportion healthy and the mean should be derived from the frequency of zeroes in the negative binomial:

$$(1-x) = \left(\frac{m}{k} + 1\right)^{-k} \quad \text{or} \quad [6]$$

$$n = Kk[(1-x)^{-1/k} - 1].$$

Table 1. Frequency distributions of lesions per leaf, leaflet, or plant compared to the Poisson and negative binomial functions

Pathogen	Host	Date (1980), day/month	Distributions	Observations, no.*	Not Poisson†	Not negative binomial†	$k^\ddagger$
<i>Puccinia peridermiospora</i>	<i>Fraxinus americana</i>	11–29/VI	11	24–235 ^t	9	0	0.45–4.58
<i>Venturia inaequalis</i>	<i>Malus pumila</i>	20/V–5/VI	7	40–311 ^e	7	0	0.18–3.82
<i>Puccinia sorghi</i>	<i>Zea mays</i>	15–23/VIII [§]	35	10–15 ^f	28	0	0.13–2.26
<i>Gymnosporangium clavipes</i>	<i>Amelanchier arborea</i>	7–15/VI	3	92–118 ^e	3	1	0.40–0.52
<i>G. juniperi-virginianae</i>	<i>Malus pumila</i>	15–18/VIII	8	44–88 ^e	8	0	0.12–0.86
<i>Phytophthora malorum</i>	<i>Malus pumila</i>	26/V–8/VI	18	42–294 ^e	14	1	0.08–1.43
<i>Pseudomonas lachrymans</i>	<i>Cucumis sativus</i>	15/VIII–5/IX	11	113–132 ^e	11	0	0.01–0.11
<i>Xanthomonas vesicatoria</i>	<i>Capsicum frutescens</i>	12/VIII–16/IX	15	64–123 ^e	14	1	0.01–0.33
<i>Dibotryon morbosum</i>	<i>Prunus serotina</i>	26/VI	1	50 ^p	1	0	0.45
<i>Peronospora tabacina</i>	<i>Nicotiana tabacum</i>	12–13/VII [¶]	3	261–555 ^p	3	2	2.40
Total			112	10–555	98	5	0.01–4.58

* t, Leaflets; e, leaves; f, five-leaf plant; p, plants.

† Distributions not fitted as shown by χ^2 , U , or T .

‡ Range of maximum likelihood estimates of the parameter k of the negative binomial function fitted to distributions not fitted by Poisson and fitted by negative binomial.

§ 1978.

¶ 1955.

FIG. 2. The fraction $(1 - x)$ of the leaves, leaflets, or plants healthy as a function of the mean number of lesions. Curves show the expectation calculated by the negative binomial function for $k = 0.25, 1$, and 2 and by the Poisson (P). Distributions of lesions are fitted by the negative binomial but not by the Poisson and had healthy hosts. Some points in upper left are obscured by crowding.

Curves in Fig. 2 are drawn for k of $0.25, 1$, and 2 . The 86 frequency distributions represented in Fig. 2 had some healthy hosts, were not fitted by the Poisson, and were fitted by the negative binomial.

When introducing the multiple infection transformation, Gregory (1) was aware of an error from over-dispersion and drew a figure much like Fig. 2. Examining observations available to Gregory, we found a median k of 0.3 for cedar-apple rust (12) and 0.3 for white pine blister rust (13, 14).

In rust and mildew of wheat (15) and in 14 of our 112 cases, the Poisson and Eq. 2 will serve. In many cases, however, Eq. 2 inaccurately estimates lesions or propagules from the proportion of diseased hosts. Many points in Fig. 2 indicate that Eq. 6 with $k = 1$ would be more accurate. The ratios between n calculated for $k = 1$ and n estimated by Eq. 2 are 1.01 and 1.05 at $x = 1\%$ and 10% but increase to 1.44 and 2.50 at $x = 50\%$ and 80% . At $x = 99\%$, the ratio is fully 21 . That is, more propagules are required to cause severe disease if dissemination is not random and $k = 1$.

The growth law can be written

$$dN/dt = dn/dt \, dN/dn \quad [7]$$

In the logistic law (Eq. 3 and curve 3 in Fig. 1) the rate dn/dt of production of propagules is rN or proportional to the number of infected hosts, and their rate dN/dn of infecting healthy hosts is the healthy proportion $(K - N)/K$. The effect of over-dispersion upon success of propagules infecting healthy hosts has been incorporated (15) in the growth law by deriving dN/dn from Eq. 6 and retaining dn/dt equal to rN :

$$dN/dt = rN(1 - x)^{(1+1/k)} \quad [8]$$

If k is 1 , curve 8 (1) shows that the increase of diseased hosts slows noticeably when a quarter of them are diseased. Compared to random distribution and the same production dn/dt of propagules, contagion causes more multiple infections and slower increases in N and x .

The rate dn/dt of production of propagules is subject to modification because the number of lesions per host often increases as the proportion x of diseased hosts increases. If the number of lesions is $-K \ln(1 - x)$ as Eq. 2 predicts for random distri-

bution and if dn/dt is proportional to the lesions rather than diseased hosts, then the growth law is

$$dN/dt = -rK[\ln(1 - x)](1 - x). \quad [9]$$

Relative to the logistic law with the same r , this accelerates the increase in diseased hosts because the number of lesions increases faster than the number of diseased hosts, and curve 9 in Fig. 1 curves upward from the straight curve 3 produced by the logistic law.

The growth law for contagious distribution, Eq. 8, can also be modified for production of propagules proportional to lesions. Using Eq. 6 to estimate lesions, one obtains

$$dN/dt = rKk[(1 - x)^{-1/k} - 1](1 - x)^{(1+1/k)} \quad [10]$$

Again, production of propagules proportional to lesions rather than diseased hosts accelerates the increase in N . The curves 10 (0.1) and 10 (1), representing Eq. 10 for $k = 1/10$ and 1 , rise more rapidly than the corresponding curves representing Eq. 8 (Fig. 1).

Finally, how great is the discrepancy between the logistic law and the modifications required because the distribution of a range of pathogens is contagious rather than random? One of us (16) has pointed out that changing the dn/dn hosts infected per propagule to conform to the negative binomial produced Eq. 8 and, when $k = 1$, a downward curve on coordinates of logit and time. Nevertheless, many contagious diseases increase on a straight line (2). Although influences of periods of incubation and infectivity can be invoked, the arguments are uncertain and the paradox remains that, although the distributions of many diseases are contagious, the logits of disease increase linearly.

A resolution of this paradox lies in modifying both dn/dt and dN/dt to produce Eq. 10. When the production of propagules is proportional to the number of lesions, when the number of hosts infected per propagule is derived from the negative binomial, and when $k = 1$, Eq. 10 reduces to the classic logistic law, Eq. 3, and a straight line on logit-by-time coordinates. This is why Eq. 10 with $k = 1$ produces a straight line on Fig. 1. This probably is why many contagious diseases that cannot be represented by the Poisson function or the multiple infection transformation nevertheless increase steadily in logits.

The authors thank C. Leben (Ohio State University) and R. E. Stall (University of Florida) for supplying cultures of bacterial plant pathogens and F. J. Anscombe and G. E. Hutchinson of Yale University for critically reviewing the manuscript.

- Gregory, P. H. (1948) *Ann. Appl. Biol.* **35**, 412-417.
- van der Plank, J. E. (1963) *Plant Diseases: Epidemics and Control* (Academic, New York).
- Anscombe, F. J. (1950) *Biometrika* **37**, 358-382.
- Bliss, C. I. & Fisher, R. A. (1953) *Biometrics* **9**, 176-200.
- Rouse, D. I., Nelson, R. R. & MacKenzie, D. R. (1980) *Phytopathology* **70**, 951-954.
- Student (1919) *Biometrika* **12**, 211-215.
- Campbell, C. L. & Pennypacker, S. P. (1980) *Phytopathology* **70**, 521-525.
- Byrde, J. W., Clark, G. M. & Welford, P. J. (1973) *Trans. Br. Mycol. Soc.* **60**, 580-583.
- Strandberg, J. (1973) *Phytopathology* **63**, 998-1003.
- Rich, S. & Waggoner, P. E. (1979) *Plant Dis. Rep.* **63**, 1012-1015.
- Waggoner, P. E. & Taylor, G. S. (1955) *Plant Dis. Rep.* **39**, 79-85.
- Schneiderhahn, F. J. (1926) *Va. Agric. Exp. Sta. Bull.* **245**, 1-35.
- Fracker, S. B. (1936) *J. Econ. Entomol.* **29**, 923-940.
- Childs, T. W. & Kimmey, J. W. (1938) *J. Agric. Res.* **57**, 557-568.
- James, W. C. & Shih, C. S. (1973) *Phytopathology* **63**, 183-187.
- Waggoner, P. E. (1981) *BioScience* **31**, 315-319.