Group analysis using ELISA : determination of the level of transmission of Soybean Mosaic Virus in soybean seed

Yves Maury, Camille Duby, Jean-Marie Bossennec, Geneviève Boudazin

Group analysis using ELISA : determination of the level of transmission of Soybean Mosaic Virus in soybean seed

Yves MAURY, Camille DUBY (*), Jean-Marie BOSSENNEC & Geneviève BOUDAZIN

I.N.R.A., Station de Pathologie végétale, F 78000 Versailles
(*) *I.N.A.-P.G., Chaire de Mathématiques, 16, rue Claude-Bernard, F 75231 Paris*

SUMMARY

The work on group testing described here is complementary to experiments reported previously (MAURY *et al.,* 1983) dealing with the potential of ELISA in testing soybean seed for soybean mosaic virus. The biological part of this paper (part I) shows that group testing gave similar results when the determination of the percentage of transmission was done using, in a comparative way, groups of 30 embryos, groups of 30 axes or groups of 30 seeds. Whatever the number of infected testas in the groups of 30 seeds, it was possible to prevent the viral antigen from testas from altering the results : indeed, after soaking seeds and grinding them briefly in a blendor, it was observed that only the embryos were ground. These results show that the procedure is practical for routine use. Part II gives mathematical elements allowing an estimate of the percentage of transmission with confidence intervals. Charts enable one to read these values directly for n = 30 seeds per group and N = 10, 30, 60 and 200 groups. An aim of this part was also to guide a good choice of n and N and to plan analyses. These mathematical data could potentially be used for group analysis with any other biological material.

Additional key words : *Charts, confidence intervals.*

RÉSUMÉ

Utilisation de la technique ELISA pour une analyse par groupes : détermination du taux de transmission du virus de la mosaïque du soja par la semence de soja.

Le travail présenté dans cette publication est complémentaire d'expériences rapportées précédemment (MAURY *et al.,* 1983) sur les potentialités de la technique ELISA pour évaluer l'infection de la semence de soja par le virus de la mosaïque du soja.
Dans la partie I, il est montré que l'analyse par groupes donne, dans les conditions décrites, des résultats similaires selon que la détermination du pourcentage de transmission est effectuée en utilisant, de façon comparée, des groupes de 30 embryons, des groupes de 30 axes embryonnaires ou des groupes de 30 graines. L'antigène des téguments, qui n'est pas impliqué dans le phénomène de transmission par la graine, ne doit pas être pris en compte dans le test. La phase d'obtention des embryons avant le test est rendue inutile par l'observation suivante : lors d'une brève homogénéisation de graines préalablement gonflées, les embryons, plus friables, sont broyés différentiellement ; les téguments, à peine lacérés, ne libèrent pas leur antigène dans le tampon d'extraction. Cette observation rend le test utilisable dans la routine.
La partie II de cette publication donne des éléments mathématiques nécessaires à l'estimation du pourcentage de transmission dans la population de graines en fonction du nombre de groupes qui réagissent positivement en ELISA.
Ce résultat peut être directement lu sur des abaques qui portent aussi les intervalles de confiance aux niveaux 0,90 ; 0,95 ; 0,99 ; 0,999. A titre d'exemple sont tracés les abaques pour n = 30 graines par groupe et N = 10, 30, 60 et 200 groupes qui se rapportent au système virus de la mosaïque du soja/semence de soja cv. « Altona ». Les abaques pour des valeurs quelconques de n et N sont aussi disponibles pour des applications de l'analyse par groupes à tout autre matériel biologique. Des conseils d'utilisation de cette méthode sont donnés ; en particulier, le lecteur est guidé pour le choix de n et de N en fonction de la précision recherchée et pour la planification de son analyse.

Mots clés additionnels : *Abaques, intervalles de confiance.*

GENERAL INTRODUCTION

The percentage of transmission of soybean mosaic virus (SMV) in a batch of soybean seed can be determined using biological or serological methods. It was previously shown that ELISA is suitable for such determinations : indeed SMV has already been easily detected in the embryos of infected seeds (BOSSENNEC & MAURY, 1978 ; LISTER, 1978 ; MA DEFANG *et al.*, 1982).

For *high percentages of transmission,* the number of seeds to be examined to obtain an accurate result is low. In such cases, individual embryo tests can easily be carried out.

However, for *low percentages of transmission,* the number of seeds to be examined is such that seed-by-seed testing would be impracticable. A method of determining such transmission rates consists of dividing the lot to be analysed into N groups of n seeds each. So N tests are then done, instead of $N \times n$ when using single embryo tests. Elsewhere, this method has widely been recommended for several virus transmission problems (GIBBS & GOWER, 1960 ; PELET, 1965 ; MARROU & MESSIAEN, 1967 ; RUSSO, 1982 ; FALK & PURCIFULL, 1983).

For each seed/virus system, two problems arise for defining the *group* : 1. its size ; 2. its composition.

1) The maximum size of a group of seeds compatible with the detection of one infected embryo in this group has to be determined to exploit this method in the most economical way.

To answer this question, we previously adopted (MAURY *et al.*, 1983) an analytical approach : we studied how SMV was distributed in a population of infected soybean seeds. Each seed was dissected into testas and embryo, each embryo into axis and cotyledons. Each part was then ground in a volume of buffer corresponding to the same weight/volume ratio and tested using ELISA to find the dilution limit compatible with the detection of SMV in *the least infected embryo* in this population. The dilution limit was found to be 1/200 w/v, i.e. for cv. 'Altona', 1 embryo in 30 ml buffer.

Various numbers of embryos were ground in 30 ml buffer. The resulting puree thickened with increasing number of embryos, making it difficult to obtain a supernatant for ELISA tests when more than 30 embryos were ground in this buffer volume. We thus deduced from this analysis that *30 seeds* is the *maximum size of a group of seeds* compatible with the detection of one infected embryo. Actually, this value does not take into account some cases where only the axis is infected and where the virus titer in this axis is very low.

2) One complication in detecting SMV in groups of seeds is the presence of virus antigen in testas. The preceding analysis showed that the concentration of antigen in testas is low but significant. The virus in the testas of soybean seeds does not contribute to infection of the seedling. We have therefore inquired whether, in attempting to determine the percentage of SMV seed-transmission in a seed batch, it is possible to ELISA-test whole seeds or instead necessary to develop a technique for removing testas before seed testing or for mechanically extracting embryo axes.

From our preceding analysis, it was concluded that testas from one infected seed added to a group of 30 healthy embryos would not contribute enough antigen for detection by ELISA, but that twice to three times this level would produce a false positive reaction, i.e. virus detection unrelated to seed-transmitted virus.

Moreover, if we consider seed lots of different varieties having each, for example 1 % infected embryos, the percentage of infected testas will differ between varieties : the lower the varietal incidence of transmission the higher the percentage of infected testas. Therefore, the lower the varietal incidence of transmission the higher the probability of having, in a group of 30 seeds, zero infected embryos and at least two infected testas. It is thus necessary to study for different varieties the consequence of ELISA-testing whole seeds instead of embryos.

The aim of this paper, part I, is thus to define the composition of the group to be used in ELISA and to confirm our preceding results through :

1) comparative tests for the determination of the percentage of transmission of a batch of seed (cv. 'Altona') using successively groups of *30 embryos,* groups of *30 axes* and groups of *30 seeds* ;

2) comparative tests using groups of *embryos* and groups of *seeds* when the varietal incidence of transmission is artificially modified from 30 % to 5 %.

With the biological parameters defined, to make the test operational it is necessary to choose, for a given determination, the optimal number n of seeds in each group (n is not necessarily the maximum number) and the number N of groups to be tested using ELISA.

The aim of this paper (part II), is to give some statistical data (arguments and graphs) and possible strategies for optimizing the use of the group analysis method.

I. COMPOSITION OF THE GROUP FOR GROUP TESTING USING ELISA

A. Definitions, materials and methods

The *transmission rate* (or transmission level) of SMV in a batch of soybean seed corresponds to the ratio : number of infected embryos/number of seeds considered (BOSSENNEC & MAURY, 1978). We will symbolise it by p. The *percentage of transmission* is therefore $100 \times p$.

When the seeds are collected only from early infected plants, p values reach a maximum value characteristic of individual cultivars : we refer to this varietal character as *"varietal incidence of transmission"* (VIT).

Details concerning seeds of cv. "Altona" and the ELISA test were given in a preceding paper (MAURY *et al.*, 1983). Microplates were Linbro microplates quality EIA. ELISA optical densities (O.D. 405 nm) were recorded using a Titertek Microelisa Reader.

Due to the lack of a constant internal SMV-concentration standard, the results of experiments not performed on the same day were not pooled but given separately for each microplate.

Each sample was incubated in duplicate wells. The O.D. 405 nm value of each sample is the mean value. Healthy controls had to be used for defining the healthy O.D. range : 8 and 10 healthy samples were used, according to the experiments, on each microplate. If the O.D. distribution of the different healthy samples is considered as being normal and if we select the coefficient in a "t distribution" table (Student Fisher) by which the standard deviation (s) has to be multiplied to give a value which has a 0.005 probability to be exceeded, then the healthy limit on a given microplate is :

$\bar{x}$ + 3.25 s when referring to 10 healthy values
$\bar{x}$ + 3.5 s when referring to 8 healthy values

($\bar{x}$ being the average value).

B. Results

1. *Comparative determination of the percentage of transmission of SMV in a seed batch using successively, groups of embryos, groups of axes and groups of seeds*

(This order for giving the results corresponds to an order of increased technical facility for obtaining large numbers of embryos, axes and seeds).

The batch of seed consisted of a mixture of seeds collected from healthy and infected soybean "Altona" in such a way that the percentage of transmission be between 1 and 2 %.

a) *ELISA test with 60 groups of 30 embryos*

— Extraction of embryos

After a 20-h soaking period at 20 to 25 °C, testas were removed by manual dissection. The embryos were distributed into groups of 30 embryos. Each group was ground in 30 ml PBST-PVP using a Waring blendor. A part of the resulting suspension was poured through fiber glass into a test tube and, after decantation, the supernatant incubated in duplicate wells. 16 groups of 30 healthy embryos were prepared in the same way. Duplicate microplates (1a and 1b) were used, each with supernatants from 8 healthy embryo groups and 30 embryo groups to be tested.
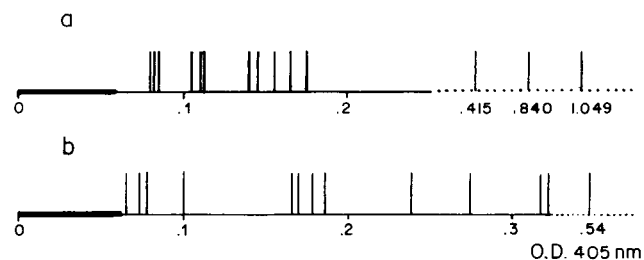

Figure 1

ELISA positive values for 60 groups of 30 embryos. Each line corresponds to one microplate. The healthy zone is represented by a thick segment on the line and the positive values by vertical bars.

Valeurs positives d'un test portant sur 60 groupes de 30 embryons en ELISA. Chaque ligne correspond à une plaque. La zone des valeurs de référence pour les groupes sains est représentée sur la ligne par un segment épais ; les valeurs pour des groupes contenant des embryons infectés par des barres verticales.

— O.D. values for healthy groups of embryos

For plate 1a, the average O.D. value for healthy groups of embryos was $\bar{x}$ = 0.032 and the standard deviation s = 0.007. All healthy groups gave values lower than $\bar{x}$ + 3.5 s = 0.058. For plate 1b, $\bar{x}$ = 0.030, s = 0.009, $\bar{x}$ + 3.5 s = 0.062.

— Infected samples

As shown on figure 1, 14 groups gave O.D. values higher than 0.058 on plate 1a and 13 groups gave values higher than 0.062 on plate 1b. So 27 groups out of 60 had some infected embryos.

This result corresponds to a percentage of infected embryos included between 1.3 and 2.8 %, at the 0.95 confidence level (see fig. 6d).

b) *ELISA test with 60 groups of 30 axes*

— Extraction of the axes

Large numbers of axes were prepared according to the method of BRUENING & RUSSEL for French bean as reported by JAFARPOUR *et al.* (1979), slightly modified : dry seeds were crushed in a laboratory threshing machine for cereals (Walter & Wintersteiger, K. G. model, Austria) adjusted for splintering the seeds. The resulting mixture of cotyledons, axes and testas was then successively sieved on 3.15,2, 1.25 and 0.8 mm meshes. All intact axes were retained on this last sieve and also cotyledons and testa fragments. The contaminants floated to the surface of water when the axes swelled and sedimented.

Axes were either used immediately or dried and ground later. This method yielded more than 50 % intact embryo axes (number of axes/initial number of seeds). The population obtained in this way was representative of the initial population of seeds. (See results below.)

Groups of 30 axes were ground in 1.2 ml PBST using a pestle and mortar. After decantation, the supernatant was tested in duplicate wells.

Duplicate microplates (2a and 2b) were used, for testing 8 axis groups from healthy seeds and 30 axis groups to be tested.

— O.D. values for groups of healthy axes

Plate 2a : $\bar{x}$ = 0.162   s = 0.025
                                    Healthy limit = 0.250.
Plate 2b : $\bar{x}$ = 0.135   s = 0.025
                                    Healthy limit = 0.223.
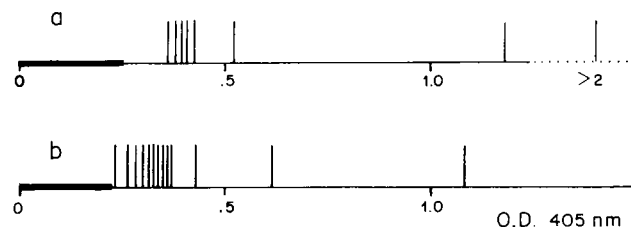

Figure 2

ELISA positive values for 60 groups of 30 axes. See fig. 1 legend.
Valeurs positives d'un test ELISA portant sur 60 groupes de 30 axes embryonnaires. Même légende que fig. 1.

— O.D. values for infected samples

As shown on figure 2, 8 groups gave O.D. values higher than 0.250 on microplate 2a and 13 groups gave values higher than 0.223 on microplate 2b. Thus, 21 groups contained one or more infected axes.

These results indicate that between 0.9 and 2.1 % of the embryo axes were infected, estimated with a 0.95 level of confidence, as deduced from figure 6d.

c) *ELISA test with 60 groups of seeds*

— Preparation of the samples

After a one night soaking period, groups of 30 seeds were ground in the same way as the preceding groups of embryos. Duplicate microplates (3a and 3b) were used, each with 10 healthy seed groups and 30 seed groups to be tested.

— O.D. values for healthy groups of seeds

Plate 3a : $\bar{x} = 0.092$, s $= 0.026$.
Plate 3b : $\bar{x} = 0.028$, s $= 0.013$.

All the healthy group had O.D. values lower than $\bar{x} + 3.25$ s ; this healthy limit was calculated to be 0.177 for plate 3a and 0.070 for plate 3b.

— O.D. values for infected samples

As shown in figure 3, 10 groups had O.D. values higher than 0.177 on microplate 3a and 11 groups had values higher than 0.070 on microplate 3b. Thus, 21 groups of seeds out of 60 had infected seeds. These results indicate that SMV would have been transmitted in between 0.9 and 2.1 % of the seeds, at a confidence level of 0.95, as deduced from figure 6d.
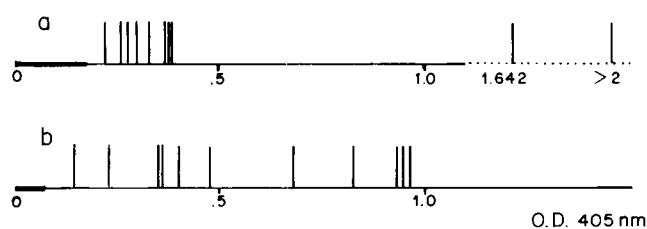

Figure 3

ELISA positive values for 60 groups of 30 seeds. See fig. 1 legend.
Valeurs positives du test ELISA pratiqué sur 60 groupes de 30 grai-
nes. Même légende que fig. 1.

d) *Conclusion*

The 3 confidence intervals at the level 0.95 are 1.3-2.8, 0.9-2.1, 0.9-2.1 respectively for embryos, axes and seeds.

These intervals share a common part, 1.3.-2.1. This means that testing 'Altona' soybean seed for SMV seed transmission using groups of 30 units gives reproducible results. Another interesting finding concerns the antigen from testas which, in the test with whole seeds, did not alter the result. These results differ with our previous deduction (MAURY *et al.,* 1983) that a group of 30 seeds having no infected embryo and at least two infected testas would presumably react positively in ELISA.

This difference could be due to :

— either the cultivar used which, owing to its high varietal incidence of transmission (30 %) and consequently its low ratio infected testas/infected embryos, would not best show that SMV antigen from testas can induce an overestimation of the percentage of transmission,

— or to the experimental conditions which could have prevented the antigen from testas from reacting in the test.

These two hypotheses were thus tested by estimating the percentage of transmission while varying the varietal incidence of transmission and the mode of grinding.

2. *Influence of infected testas on ELISA estimation of SMV transmission rate*

a) *Artificial variation of the varietal incidence of transmission (VIT)*

When mixing 27 seeds collected from early infected 'Altona' plants (about 9 infected embryos) with 873 healthy seeds collected from healthy 'Altona' plants, the resulting batch of 900 seeds has a transmission percentage close to 1 %. In this experiment, 153 among 873 healthy seeds were dissected and their healthy testas replaced by infected testas in order to increase the ratio of infected testas/infected embryos, thus mimicking a variation of the VIT from 30 to 5 %. The purpose was to test ELISA for estimating the percentage of transmission for this synthesized VIT, using 30 groups of 30 seeds.

To have a random distribution of the infected seeds and infected testas in these 30 groups we worked as shown in table 1 (left part) :

— a first random draw distributed the 27 seeds from infected plants (with about 8 or 9 infected embryos),

— a second random draw distributed 153 infected testas (in a cv. having a VIT $= 5$ %, the ratio of infected testas/infected embryos ≥ 20. If we suppose 9 infected embryos among 27 infected seeds, the number of testas to make a VIT $= 5$ % is thus $9 \times 20 = 180$. So $180 - 27 = 153$ testas were to be added at random),

— as many healthy embryos as infected testas were then added to each group to reconstitute 'mixed' seeds,

— each group was then brought to 30 with healthy seeds.

Grinding each group was performed in two steps, 30 s each at high speed in a Waring blendor. In the first step all the seed material was ground, except the infected testas that were kept separately ; a sample of 2 ml was taken. In the second step, the infected testas were added. In this way it was possible to measure for each group the amount of antigen contributed by testas. It was also possible to compare the results corresponding to VIT $= 30$ % (infected testas not added) to the results corresponding to VIT $= 5$ % (infected testas added).

Two microplates were used. On each were incubated

TABLE 1

Composition of 30 groups of 30 seeds for studying the influence of presence of infected testas on ELISA estimation of SMV transmission rate.
The seed lot considered had a transmission percentage close to 1 % and a Synthesized VIT = 5 %.
(I.S. = infected seeds, from the original batch with a 30 % transmission. I.T. = infected testas. H.E. = healthy embryos. H.S. = healthy seeds).
Constitution de 30 groupes de 30 graines pour analyser l'influence de la présence de téguments infectés sur la détermination par ELISA du taux de transmission.
Le lot considéré a un pourcentage de transmission proche de 1 p. 100 et un VIT synthétique de 5 p. 100.
(I.S. = graines infectées, du lot d'origine qui transmet le virus à environ 30 p. 100. I.T. = téguments infectés. H.E. = embryons sains. H.S. = graines saines).

| | Group composition | | | ELISA O.D. 405 nm $\times$ 10^3 | | |
				Infected testas		Control healthy groups.
n°	I.S.	I.T.	H.E.	H.S.	Not added	Added	
1	1	5	5	24	673	670	194
2	2	9	9	19	346	336	213
3		6	6	24	130	154	277
4		7	7	23	187	187	214
5		13	13	17	235	243	217
6		4	4	26	166	210	190
7		4	4	26	177	196	228
8		5	5	25	155	203	231
9	2	6	6	22	828	703	207
10	3	3	3	24	927	1 077	244
11	1	4	4	25	246	218	
12	3	3	3	24	241	287	
13		6	6	24	179	217	
14		6	6	24	167	218	
15	2	4	4	24	184	210	
16		2	2	28	193	178	140
17	3	5	5	22	345	314	144
18	1	7	7	22	141	178	129
19		8	8	22	139	166	178
20	2	5	5	23	1 190	1 256	158
21	1	5	5	24	492	588	172
22		2	2	28	150	130	146
23		6	6	24	132	163	167
24		5	5	25	125	155	170
25	1	4	4	25	236	294	181
26	1	4	4	25	158	154	
27		6	6	24	154	193	
28	2	3	3	25	144	155	
29		3	3	27	140	160	
30	2	3	3	25	948	1 029	
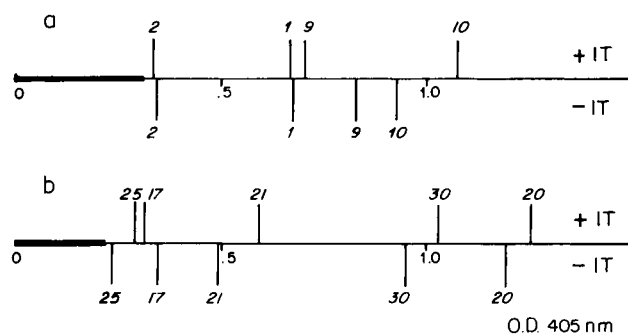
Figure 4
Influence of the presence of infected testas (IT) in groups of seeds on their ELISA value.
The number of each bar refers to the number of a group defined in table 1.
Influence de la présence de téguments infectés (IT) dans les groupes de graines sur leur réaction en ELISA.
Le numéro de chaque barre renvoie au numéro d'un groupe défini dans le tableau 1.

in duplicate wells the samples from :
— 15 groups (first step of grinding)
— the same 15 groups + the corresponding infected testas (second step)
— 10 groups of 30 healthy seeds.

O.D. 405 nm values are reported in table 1 (right part). On figure 4, for each microplate, the healthy limit and the positive reactions in presence and absence of infected testas are drawn.

It is noteworthy that even in the presence of several infected testas, the groups of seeds gave about the same ELISA values as did the corresponding groups of embryos ; particularly, group n° 5 with 0 infected embryos and 13 infected testas did not give a positive ELISA reaction.

b) *Influence of different modes of grinding*

Two groups of 20 infected testas were manually prepared after a one-night seed-soaking period. Each testa was divided into two halves using a scalpel to constitute 2 $\times$ 2 groups of 20 half testas t_a similar to $t_a(m)$ and t_b similar to $t_b(m)$. $t_a(m)$ and $t_b(m)$ were ground in 4 ml PBST-PVP using mortar and pestle.

12 groups of 30 healthy embryos were also prepared to be ground using a Waring blendor (30 ml PBST-PVP, high speed, 30 s). As shown on table 2, a sample was taken after grinding from groups n° 9 and 10 before adding t_a and $t_a(m)$ respectively and grinding again. On the other hand, t_b and $t_b(m)$ were added to groups n° 11 and 12 respectively before grinding. O.D. 405 values are reported in table 2.

TABLE 2

ELISA values of groups of healthy seeds containing some infected testas depending on the mode of grinding testas (Waring blendor a, b ; mortar a (m), b (m)).
Valeurs obtenues en ELISA pour des groupes de graines saines contenant un nombre variable de téguments infectés selon le mode de broyage des téguments (Waring blendor a, b ; mortier a (m), b(m)).

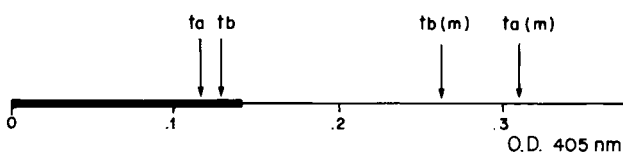| Group n° | Group composition | ELISA O.D. 405 nm |
1	30 healthy embryos	0.093
2	»	0.111
3	»	0.093
4	»	0.080
5	»	0.066
6	»	0.084
7	»	0.103
8	»	0.101
9	30 healthy embryos	0.110
9'	» + t_a	0.114
10	30 healthy embryos	0.088
10'	» + $t_a(m)$	0.310
11	30 healthy embryos + t_b	0.122
12	30 healthy embryos + $t_b(m)$	0.262

Figure 5
False positive ELISA reactions induced by carefully ground infected testas — a (m), b (m) —.
Induction de réactions faussement positives par un broyage au mortier — a (m), b (m) — des téguments infectés.

Figure 5 shows more clearly that the antigen from testas participated in the reaction only when testas were ground using mortar and pestle.

C. Discussion

The aim of the reported experiments was to explain and define the conditions of group testing using ELISA for the system soybean seed/soybean mosaic virus.

The tests performed using groups of 30 embryos in 30 ml buffer or groups of 30 axes in 1.2 ml buffer gave uniform results, thus confirming previous deductions (MAURY *et al.*, 1983). It also confirmed that the antigen from testas, when involved in the test, induced an overestimation of the percentage of transmission. A first technical approach for eliminating the testas was to develop a method of preparation of the axes from the lots to be tested. This method worked with a yield of intact axes higher than 50 %. However the easiest way to prevent the antigen in testas from participating in the test consisted of grinding the embryos selectively with a Waring blendor after having soaked the seeds overnight, rendering the embryos friable and testas resistant to blending.

The scheme for a routine test was thus very simple : sampling, grouping (30 seeds maximum in each group) ; for each group, soaking (one night, 25 °C) grinding (Waring blendor, 30 s, 30 ml Buffer) and clarifying (10 000 g, 15 min). The supernatant was then incubated on an ELISA microplate. The most appropriate strategy for determining the percentage of transmission with a given accuracy (number of seeds in each group, and number of groups to be tested) is discussed in the following part II.

Before any utilization of these results, some comments are appropriate.

First, the proposed method for the system soybean seed/SMV was actually elaborated for the system cv. 'Altona' seed/strain SMV1. To extend its value, it is necessary to confirm that the serum can be used for any strain of this virus. CHO & GOODMAN (1979) and CHEN *et al.* (1982) noted no detection failures in testing their antiserum by ELISA against 7 groups of SMV strains.

Different cultivar/strain pairs must then be analysed to determine whether the size limit for one group is the same for all pairs. Finally it is also necessary to determine whether the testas of the species or cultivar to be tested are hard enough to resist the proposed homogenization conditions.

The proposed evaluation by ELISA also requires that the zone of healthy seed reaction be well defined. It was apparent in figures 1-4 that the O.D. 405 nm values for infected seed groups were often close to the values for healthy seed limits. We accordingly propose systematic reference to O.D. values for 10 healthy groups per microplate to account for differences in polystyrene adsorption from plate to plate. The healthy limit corresponding to a 0.005 probability of being exceeded is $\bar{x}$ + 3.25 s. The use of well clarified suspensions helps to lower this limit ; particularly we noticed that polyvinylpyrrolidone (1 %) in the extraction buffer made decantation of insoluble material more efficient. However centrifugation of

extracts appeared to yield superior results. In a comparative test, extract from a single healthy-seed group was subdivided and half was treated with polyvinylpyrrolidone and decanted and half was centrifuged. The extracts were tested in the 36 inner wells of a single microplate. The results, respectively, were : $\bar{x}$ = 0.116, s = 0.026 when decanted ; and $\bar{x}$ = 0.094, s = 0.014 after centrifugation.

In these conditions, this test enables one to determine with a defined accuracy relatively low levels of transmission. For example, 60 groups of 30 seeds which would give 3 positive ELISA values would correspond to about 0.16 % transmission, the confidence interval (0.95) being 0.06 − 0.5 % (fig. 6*d*).

As this method comes to the same as working with embryos, we saw (MAURY *et al.*, 1983) that a correction coefficient, the value of which could be accurately determined when testing large numbers of single seeds, should be used. However, for low percentages of transmission, this coefficient scarcely affected the confidence interval. In any case we believe that progress in the accuracy of virus detection in seeds should result not from evaluation of this correction coefficient, for each plant cultivar and virus strain, but from improved sensitivity of the serological technique utilized, enabling detection of low virus titer in axes and obviating any need for a correction coefficient.

Improved detection sensitivity would allow an increase in the maximum number of seeds in each group, and when testing the same number of groups as previously, lower transmission percentages could be determined. In the present conditions, the determination of lower transmission percentages is theoretically possible by increasing the number of groups to be tested — e.g. 5 positive ELISA reactions for 200 groups of 30 seeds (fig. 6*f*) corresponds to a percentage of transmission between 0.04 and 0.2 % (0.95 confidence interval). An alternative approach might consist of testing germinated seeds, on the supposition that virus titer increases after germination. However we did not observe any increase in titer after an 8-day germination period, in rolled paper towels, as did LISTER (1978) with another cultivar. But, after a three-week germination period in a greenhouse, primary leaves had a noticeably higher virus titer. The subsequent increase in sensitivity would be very expensive for routine testing. However, working with larger groups would be of interest for accurate evaluations of percentages of infected plants in epidemiological studies.

II. STATISTICAL INTERPRETATION AND UTILIZATION OF GROUP ANALYSIS

A. Introduction

When a sample of seeds arrives in a laboratory for analysis, two questions may arise concerning the infection rate p of a virus.

1) One may wonder whether this sample meets certain standards. In this case, one must test whether p is higher than a certain level determined in advance.

2) One can try to estimate p, stating the precision of the estimate.

The first question was studied by SHU GENG *et al.* (1983) and is particularly interesting in the framework of quality control of seed production.

The second question was approached by GIBBS & GOWER (1960) yielding results on maximum likelihood estimate and particularly showing the existence of a bias which increases with the number of seeds in a group.

Our purpose here is to give mathematical elements allowing one to estimate *p* with confidence intervals and to guide the planning of analysis when the method described in Part I is used.

The population concerned here is the set of seeds sent to the laboratory for analysis. We do not consider here the problem of this set constitution which is in itself a sample of a larger population.

The method is basically a random creation and random draw of N groups of n seeds of this population.

For each group, one must either assume that the group is infected or healthy. If a group is infected, this means that at least one seed in the group is infected (here we mean by infected seed, a seed able to transmit the virus). If a group is healthy, it means that all seeds are healthy (i.e. none are capable of transmitting the virus).

B. Estimate of p

Let Y be the number of healthy groups among the N groups analysed.

Let X_i be the number of infected seeds in the ith group. If the N groups have been constituted and taken randomly, one can say that X_i is a random variable the distribution of which is binomial with n and p parameters.

We may write :

P_r (i^{th} group is healthy) $= P_r (X_i = 0) = (1 - p)^n = q^n$ with $q = 1 - p$.

And Y is a random variable whose distribution is binomial with N and q^n parameters.

From the analysis of the N groups and from this model, we derive the maximum likelihood estimator :

$$\hat{q}^n = \frac{Y}{N}$$

From this, we deduce an estimate of p :

$$\hat{p} = 1 - \left(\frac{Y}{N}\right)^{1/n}$$

This estimate of p is valid only if the binomial model is correct for the X_i's and Y. The model is correct only if one can ensure that the results for the seeds are mutually independent, as well as the results for the groups.

Furthermore, it is necessary that the draw of seeds to form the groups and the draw of groups give to each seed of the population the same probability of belonging to the sample and give to each group the same probability of belonging to the sample.

C. Confidence intervals of p

In order to find a confidence interval of p with 1-α level (for example 95 %), we must first establish a confidence interval of q^n.

For simplicity, we give only an approximation of confidence intervals of q^n. This approximation, the quality of which depends on N, is good enough when N $\geq$ 10.

If N is large enough, Y/N has approximately normal distribution with mean q^n and variance $\frac{q^n(1 - q^n)}{N}$. The 1-α confidence interval of q^n is :

$$\frac{Y + t_\alpha^2/2 - t_\alpha \sqrt{Y(1 - Y/N) + t_\alpha^2/4}}{N + t_\alpha^2} < q^n < \frac{Y + t_\alpha^2/2 + t_\alpha \sqrt{Y(1 - Y/N) + t_\alpha^2/4}}{N + t_\alpha^2}$$

where t_α is the value read in the table of the normal variable N(0,1) such that $Pr(Z > t_\alpha) = \alpha/2$.

The function $(1 - p)^n$ being monotone, we deduce the 1-α confidence interval of p :

$$1 - \left(\frac{Y + t_\alpha^2/2 + t_\alpha \sqrt{Y(1 - Y/N) + t_\alpha^2/4}}{N + t_\alpha^2}\right)^{1/n} < p < 1 - \left(\frac{Y + t_\alpha^2/2 - t_\alpha \sqrt{Y(1 - Y/N) + t_\alpha^2/4}}{N + t_\alpha^2}\right)^{1/n}.$$

D. Charts

The previous expressions allow one to find an estimate of p and 1-α confidence intervals. But it is easier to find these results using the chart constructed from the formulas. From the number Y of healthy groups a simple reading provides in ordinate the estimate p (dotted line) and the confidence intervals at the respective levels .90, .95, .99, .999 (continuous lines) on the chart corresponding to the values N and n of the experiment (fig. 6).

E. Guide for use of the method. Planning the analyses

It is the magnitude of the confidence intervals for a given level 1-α, which gives the precision of the method. The magnitude is determined by the four parameters N, n, Y and α. We can make the following remarks :
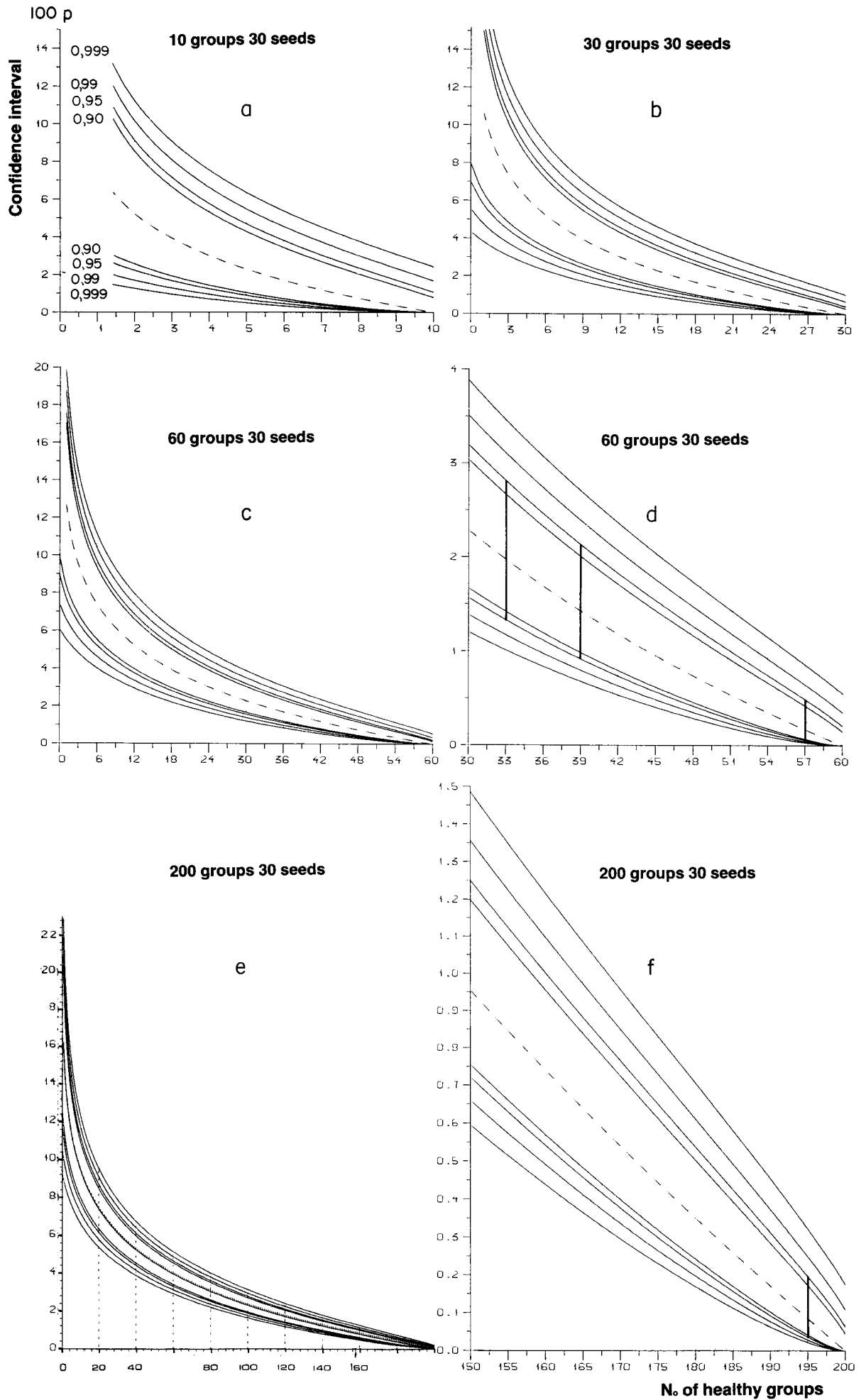
Remark 1

For N, n and 1-α given, this confidence interval increases when Y decreases, from the value

$$1 - \left(\frac{N}{N + t_\alpha^2}\right)^{1/n} \text{ for } Y = N \text{ to } \left(\frac{t_\alpha^2}{N + t_\alpha^2}\right)^{1/n} \text{ for } Y = 0.$$

This magnitude keeps small values for Y = N but can be considerable for Y = 0.

Example : for n = 50, N = 60, 1-α = .95, the magnitude of confidence intervals varies from 0.2 to 99.8 percent.

No of healthy groups

Remark 2

For n and 1-α given, and for every observed value of Y, the confidence interval magnitude always decreases with N, i.e. we always gain in precision when increasing the number of observed groups.

Remark 3

For N and α given, the magnitude of a derived confidence interval for an observed Y is not always an increasing or decreasing function of n. If Y = 0, the confidence interval magnitude $\left(\dfrac{t_\alpha^2}{N + t_\alpha^2}\right)^{1/n}$ is an increasing function of n and if Y = N, the magnitude

$1 - \left(\dfrac{N}{N + t_\alpha^2}\right)^{1/n}$ is a decreasing function of n.

This means that, if p is not very small, there is a great probability that Y = 0 and therefore there is no interest in increasing the number of seeds in the group.

However, if p is very small, there is a great probability that Y = N, therefore it is better to increase, if technically possible, the number n. The probability of Y = 0 is equal to $(1 - q^n)^N$ and the probability of Y = N is equal to q^{nN}.

Examples

For p = 5 %, n = 100, N = 10 we have :
Pr(Y = 0) = 0.94 and Pr(Y = N) = 5.10^{-23}.
For p = 0.01 %, n = 100, N = 10 we have
Pr(Y = 0) = 9.10^{-21} and Pr(Y = N) = 0.90.

If we want to control the risk that Y = 0, fixing it at the value π (for example π = 0.05) it is necessary to have :

$$N > N_{\text{limit}} = \frac{\text{Log } \pi}{\text{Log } (1 - q^n)}$$

$$n < n_{\text{limit}} = \frac{\text{Log } (1 - \pi^{1/N})}{\text{Log } q}$$

Figure 7 gives with p the values of N_{limit} for π = 0.05 and n = 30, 50 and 100.

Figure 8 gives with p the values of n_{limit} for π = 0.05 and N = 10, 20, 30, 50, 80 and 100.
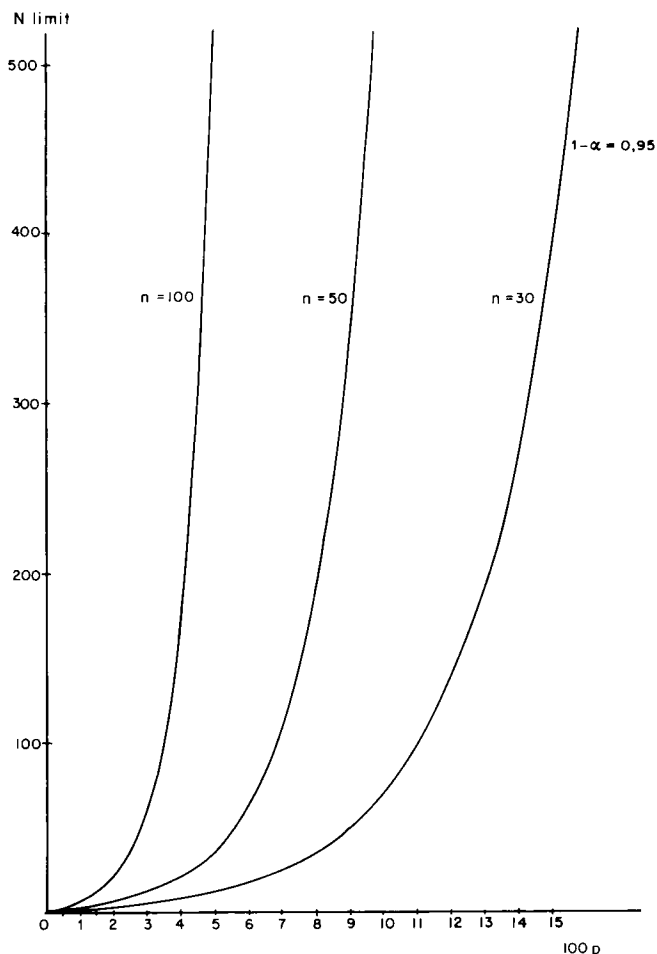

Figure 7
Curves giving as a function of the percentage of transmission (100 p) the minimum number of groups to be examined in order to control at the level 0.05 the risk of obtaining Y = 0 (for n = 30, 50, 100).
Courbes donnant en fonction du pourcentage de transmission (100 p) le nombre minimum de groupes à examiner pour contrôler à 5 p. 100 le risque d'avoir Y = 0 (pour n = 30, 50, 100).

Figure 6
Estimate of the percentage of transmission (100 p) as a function of the number of healthy groups (------) and confidence intervals at the levels 0.90, 0.95, 0.99 and 0.999 (————).
This estimate is done using N groups of 30 seeds each :
N = 10 (figure 6a), N = 30 (figure 6 b), N = 60 (figures 6c and 6d), N = 200 (figure 6e and 6f). Figures 6d and 6f are enlargements of the most useful part of figures 6c and 6e respectively.
N.B. Charts for any value of N and n are available from « INRA, Laboratoire de Biométrie, F 78000 Versailles ».
Estimation du pourcentage de transmission (100 p) en fonction du nombre de groupes sains (------) et intervalles de confiance aux niveaux 90, 95, 99 et 99,9 p. 100 (————).
Les figures 6d et 6f sont des agrandissements des parties les plus utiles des figures 6c et 6e.
N.B. Les abaques correspondant à des valeurs quelconques de N et n sont disponibles à « INRA, Laboratoire de Biométrie, F 78000 Versailles ».
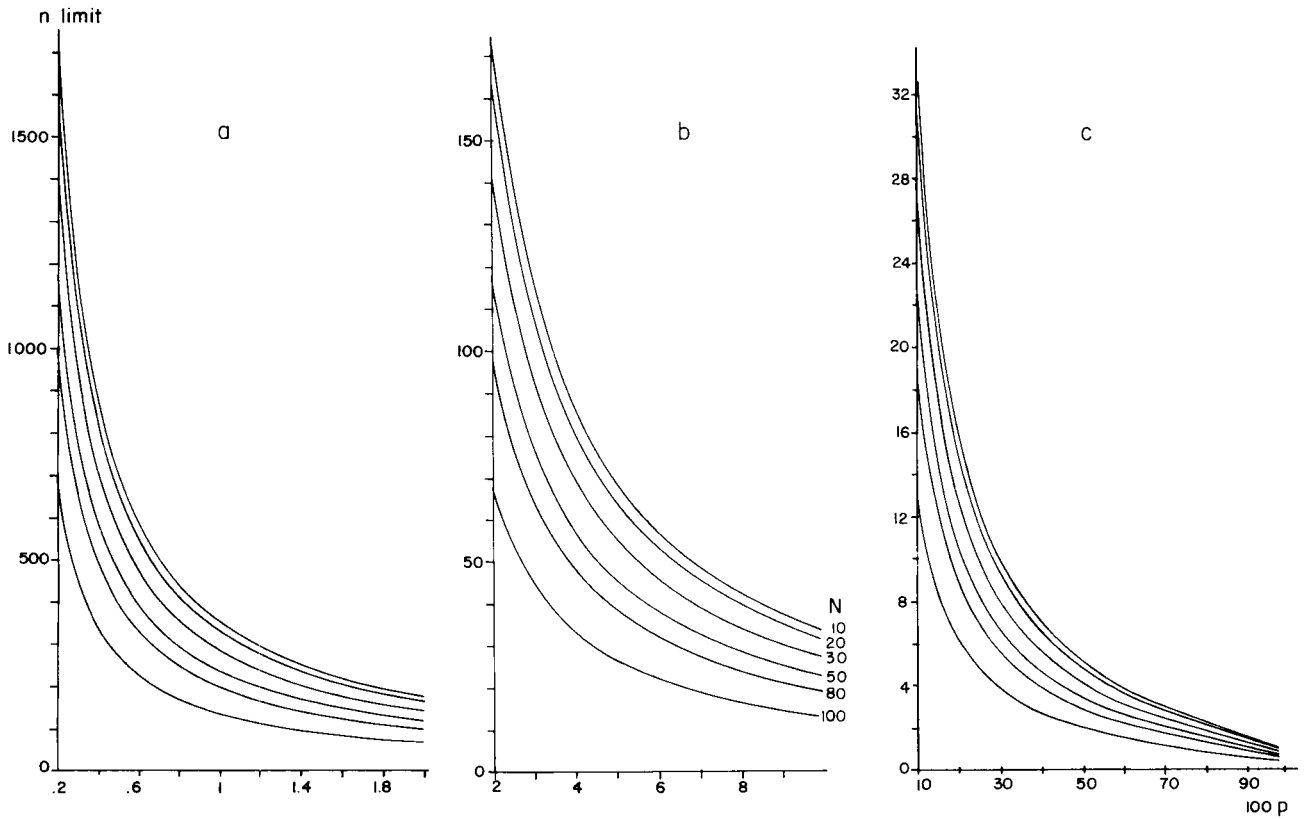
Remark 4

We can have an idea of the precision expected before the experiment, by calculating the expected magnitude of the confidence intervals. This magnitude is, of course, a function of q = 1 - p. Using a first order development, an approximation of the expected magnitude of the 1-α confidence interval of p is :

$$\Delta p = \left(\frac{Nq^n + t_\alpha^2/2 + t_\alpha \sqrt{Nq^n(1-q^n) + t_\alpha^2/4}}{N + t_\alpha^2}\right)^{1/n} - \left(\frac{Nq^n + t_\alpha^2/2 - t_\alpha \sqrt{Nq^n(1-q^n) + t_\alpha^2/4}}{N + t_\alpha^2}\right)^{1/n}.$$

We can use this last expression numerically only when p is small (less than 30 %) because the functions

n^{th} power and $\left(\dfrac{1}{n}\right)^{\text{th}}$ power are not stable.

Figure 8

Curve giving as a function of the percentage of transmission (100 p) the group size limit (n limit) in order to control at the level 0.05 the risk of obtaining Y = 0 (for N = 10, 20, 30, 50, 80 and 100). 8a : 0.2 < 100 p < 2 ; 8b : 2 < 100 p < 10 ; 8c : 10 < 100 p < 90.

Courbes donnant en fonction du pourcentage de transmission (100 p) la taille maximum des groupes (n limite) pour contrôler à 5 p. 100 le risque d'avoir Y = 0 (pour N = 10, 20, 30, 50, 80 et 100).

If p is greater than 30 %, it is better to use the following expression for Δp :

$$\Delta p = \left(\frac{Nq^n + t_\alpha^2/2 + t_\alpha \sqrt{Nq^n(1-q^n) + t_\alpha^2/4}}{N + t_\alpha^2}\right)^{1/n} - q^2\left(\frac{N}{t_\alpha^2}\right)^{1/n}.$$

The second term of this expression is a second order development in q^n and gives a better precision, because the difference

$$Nq^n + t_\alpha^2/2 - t_\alpha \sqrt{\frac{Nq^n(1-q^n) + t_\alpha^2/4}{N + t_\alpha^2}}$$

can be inferior to the precision of the calculator.

We can find an expected relative precision with $\dfrac{\Delta p}{p}$.

Figure 9 gives this relative precision for n = 30, $1 - \alpha = 0.95$, in function of N and allows one to value the cost of a given level of precision.

It should be noted that, for low infection rates, an excessive demand on the relative precision can cost considerably if n is not very high. For example, if p = 1 %, in order to obtain a relative precision lower than 1, the cost expressed in number of analyses increases very rapidly.

From the above remarks, we can consider two strategies for carrying out seed testing.
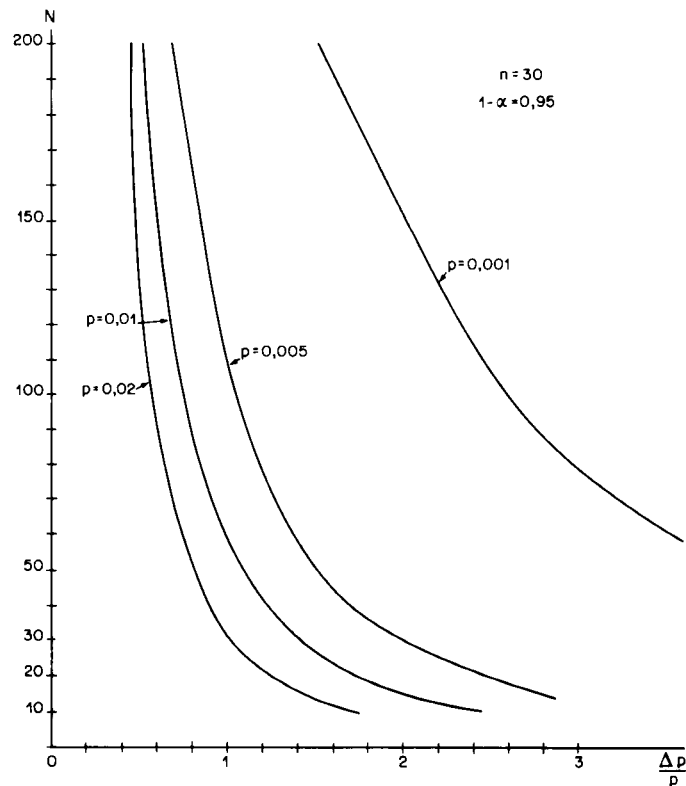

Figure 9

Curves showing for different values of p and for n = 30 the relative mean precision as a function of the number N of groups tested. Courbes donnant pour différentes valeurs de p et pour n = 30 la précision relative moyenne en fonction du nombre N d'analyses.

a) *First strategy*

We have no idea of the order of magnitude of the infection rate p and can proceed with more than one step :

a) First step : we make ten groups containing the maximum number n_{max} of seeds permitted technically by the method.

— If the result of the ten analyses gives Y ≠ 0, this means that n_{max} is a number of seeds per group well adapted to the value of p. We obtain from this Y a first estimate of p. Using figure 9 we find the number N of groups necessary to obtain a given $\dfrac{\Delta p}{p}$. We then go to the second step.

— If the result of the ten analyses is Y = 0, this means that n_{max} is too high for the value of p. We must therefore start again the first step with $n < n_{max}$ in order to obtain a value of Y different from zero. We are therefore in the previous situation.

b) Second step : We analyse the N groups of n seeds, where N and n have been determined at the first step. If $n = n_{max}$, we estimate p from a number of groups equal to N + 10.

If $n < n_{max}$, we estimate p from the result of the N analyses of the second step, forgetting the ten analyses of the first step (it is very complicated to give an estimate using the first ten analyses because the number of seeds per group is different).

b) *Second strategy*

We can have an idea of the order of magnitude of p, i.e. we know already whether we are dealing with a low, medium or high infection.

We then will use figure 7 and figure 8 in order to control the risk of a bad choice of N and n (that would give, with a high probability, Y = 0). Keeping in mind the technological constraints, the cost of analysis and the knowledge about p, we determine then a value of N and n. We will achieve the N

analyses which give an estimate and a confidence interval of p.

With this second strategy we cannot control the precision of the estimate of p, but we manage things so that the precision is not too bad.

It is better to choose the first strategy if possible and if control of the precision for the estimate of p is required.

F. Discussion

The method here presented allows us to find an estimate and a confidence interval of the seed transmission rate for a virus in a population of seeds. It also provides necessary analyses to obtain a given level of precision. The difficulty of the method proceeds from the fact that two numbers must be determined : N number of analyses and n number of seeds per group. The quality of the results depends on the unknown value p. Then, its efficiency is better when we arrange procedural strategies, the first one deriving appropriate values for both n and N.

We have avoided approximation of the binomial by means of a Poisson distribution, because this approximation is valid only when the infection rate p is very small and n large, which is not always the case. Indeed, if one wants to study the evolution of a disease epidemic, values of p can be large and this approximation would therefore be inappropriate. The results presented in figures 6-9 are valid for every value of p.

REFERENCES

Bossennec J. M., Maury Y., 1978. Use of the ELISA technique for the detection of Soybean Mosaic Virus in Soybean seeds. *Ann. Phytopathol.,* 10, 263-268.

Chen L. C., Durand D. P., Hill J. H., 1982. Detection of Soybean Mosaic Virus pathogenic strains by enzyme-linked immunosorbent assay using polystyrene plates and beads as the solid phase. *Phytopathology,* 72, 1177-1181.

Cho E. K., Goodman R. M., 1979. Strains of Soybean Mosaic Virus : Classification based on virulence in resistant soybean cultivars., *Phytopathology,* 69, 467-470.

Falk B. W., Purcifull D. E., 1983. Development and application of an enzyme-linked immunosorbent assay (ELISA) test to index lettuce seeds for lettuce mosaic virus in Florida. *Plant Dis.,* 67, 413-416.

Gibbs A. J., Gower J. C., 1960. The use of a multiple transfer method in plant virus transmission studies — some statistical points arising in the analysis of results. *Ann. Appl. Biol.,* 48, 75-83.

Jafarpour B., Shepherd R. J., Grogan R. G., 1979. Serologic detection of bean common mosaic and lettuce mosaic viruses in seed. *Phytopathology,* 69, 1125-1129.

Lister R. M., 1978. Application of the enzyme-linked immunosor-

bent assay for detecting viruses in soybean seed and plants. *Phytopathology,* 68, 1393-1400.

Ma Defang, Xu Shaohua, Zhang Zuofang, Hu Weizhen, Zhang Chengliang, Song Shumin, Zhao Ruisheng, 1982. Detection of Soybean mosaic virus with ELISA. *Acta Phytopathol. Sin.,* 12, 47-52.

Marrou J., Messiaen C. M., 1967. The *Chenopodium quinoa* test : a critical method for detecting seed transmission of Lettuce Mosaic Virus. *Proc. Int. Seed Testing Assoc.,* 32, 49-57.

Maury Y. Bossennec J. M., Boudazin G., Duby C., 1983. The potential of ELISA in testing soybean seed for soybean mosaic virus. *Seed Sci. Technol.,* 11, 491-503.

Pelet F., 1965. Dosage du virus de la Mosaïque de la Laitue par indexage de la graine sur *Chenopodium quinoa* Willd. *Rev. Hortic. Suisse,* 38, 7-10.

Russo M., 1982. Moderne metodiche per il rilevamento delle infezioni virali nei semi. II. Immunomicroscopia elettronica. *La difesa delle piante,* 5-6, 395-404.

Shu Geng, Campbell R. N., Carter M., Hills F. J., 1983. Quality control programs for seed borne pathogens. *Plant Dis.,* 67, 236-242.