

Clinical Trial of Hyposensitisation in Hay Fever: Two Methods of Relating Symptom Scores to Daily Pollen Counts

P. WILKINSON and E. TAUDORF

Clinical Research Dept., Dome/Hollister-Stier, Stoke Poges, U.K. and Allergy Unit, Medical Dept. TA, Rigshospital, Copenhagen, Denmark

Two methods of relating daily pollen counts to patients' symptom scores in clinical trials of hyposensitisation in hay fever are described. The problems involved in comparing results from different centres and years of a trial are discussed with particular reference to symptoms which are not pollen related. The Normalised Regression method is a powerful technique, but often the Pollen Interval method is preferable on the grounds of simplicity of communication and calculation.

Key words: clinical trials; hay fever; immunotherapy; pollen allergy; statistics.

Accepted for publication 19 February 1984

CLINICAL ASPECTS

There is a great need for systematic, accurate and reproducible methods for clinical evaluation of immunotherapy. Two statistical methods of relating daily pollen counts to patients' symptom scores in clinical trials of hyposensitisation in hay fever are described. It is stressed under which circumstances symptom and medication scores can be used in the clinical evaluation. The problems involved in comparing results from different centres and years of a trial are discussed with particular reference to symptoms which are not pollen related.

The importance of recording daily pollen count in obtaining the clinical history of allergic patients and in the evaluation of the treatment of respiratory allergic diseases has been emphasized (5, 10, 11, 12). A good correlation between the patient's symptoms and the pollen count has been found in a number of studies (9, 10, 16). To prove whether a new hyposensitisation treatment is superior to placebo it is common to apply many *in vivo* tests: nasal provocation test, conjunctival provocation test, bronchial provocation test, skin prick test; and *in vitro* tests: allergen-specific IgE, allergen-specific IgG, CRIE and histamine release in basophils. This is indeed of interest if these parameters show a change during therapy in favour of the new treatment. But from a clin-

ical point of view this is only really interesting if the treated patients also experience a reduction in allergic symptoms during their season. Subjective allergic symptoms recorded by the patient are so-called soft data. Although soft data they are nevertheless crucial to the final proof of whether or not a new treatment works. The daily symptom score is influenced by the amount of pollen in the air and by the anti-allergic medication. To get reliable daily symptom scores, daily medication should be recorded in a standardized manner.

In the past, relationships between symptom scores recorded by patients, and pollen counts have either been tested by calculating a correlation coefficient, or examined by graphical methods. These approaches are inadequate if

results from 2 years of a trial are to be directly compared, particularly when patients are living in areas with different pollen counts.

The effects of treatment on the occurrence of severe symptoms at different pollen levels have been examined (16). However, both clinical history and skin tests indicate that many patients suffer from multiple allergies (6, 9), and thus it is vital to analyse only symptoms caused by the allergen under study.

This paper suggests two possible quantitative approaches to these problems.

Method 1 – Normalised Regression

This method evolved from early attempts to quantify in each patient the relationship be-

tween symptom scores obtained from diaries completed daily by the patient, and the daily pollen count obtained from a pollen trap in the vicinity of the patient's home.

For each patient, a linear regression was performed of daily symptom scores against pollen count, shown theoretically as Fig. 1. However, from our experience the relationship between symptoms and pollen counts in most patients tends to follow a curve as shown in Fig. 2, instead of a straight line.

If linear regression had been applied to the curve shown in Fig. 2 the dotted line [1] would have resulted. However, if a patient received a low level of pollen challenge in a pollen season, then only the lower part of the curve would be relevant. In this case a linear regression would result in the dotted line [2] as shown. Thus, the regression line obtained depends on whether

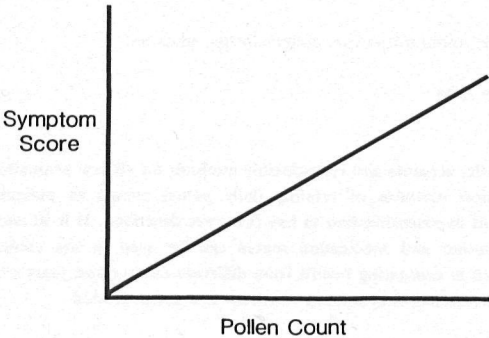


Fig. 1. Hypothetical regression of symptom score against daily pollen count.

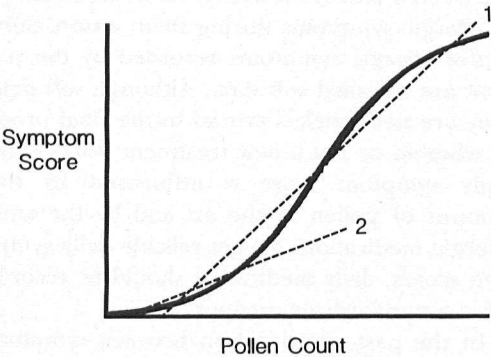


Fig. 2. A curve fits the data better than the straight line in Fig. 1. Regression line 1 would have been fitted to the complete data. Regression line 2 applies to low pollen counts only.

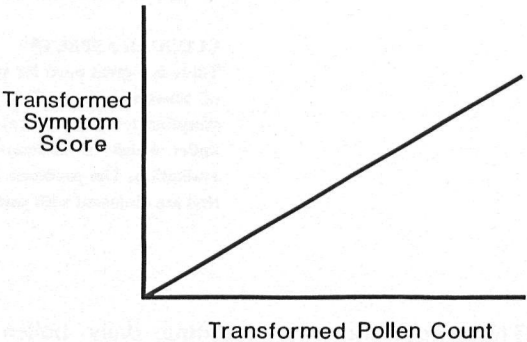


Fig. 3. Result of applying Inverse Normal Scores transformation to the curve in Fig. 2.

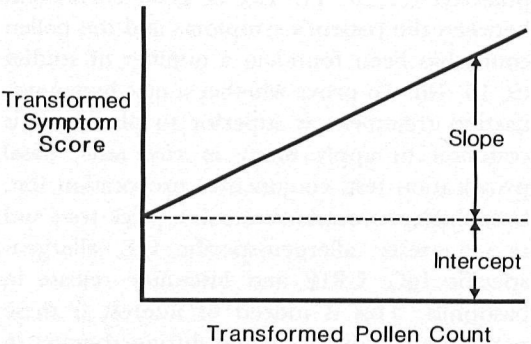


Fig. 4. More realistic graph, showing slope and intercept.

pollen levels are high or low in a particular season. This is clearly unsatisfactory.

A method which overcomes this problem requires a mathematical transformation of both symptom and pollen count data. This technique called Inverse Normal Scores (2), guarantees to remove any curvature from a graph. So long as the graphs from different patients do not differ widely in shape, each will produce a linear response as shown in Fig. 3. In practice whenever this technique has been used (3, 7, 8, 13), many graphs like Fig. 4 result.

The symptoms of each patient during a pollen season can therefore be summarised by two numbers: – the slope and the intercept of a regression line (Fig. 4). The slope represents the patient's rate of symptomatic response to increasing daily pollen challenge. The intercept represents all other symptoms, which may or may not be allergic in origin. The slopes obtained from each patient can then be analysed using standard statistical techniques, for example, Student's *t* test, analysis of variance or covariance, or Mann-Whitney U test (1, 2, 14). Thus the effects of treatment and other factors upon the severity of the disease can be assessed.

Method 2 – Pollen Count Intervals

This method involves the division of the range of daily pollen counts over one (or more) complete pollen season(s) into a number of intervals. To avoid the possibility of biasing results, the intervals should be defined before the patient diaries have been examined. In a trial involving only one pollen count, a useful method is to divide daily pollen counts as closely as possible into equal groups, for example, quartiles, so that each of the four intervals contains an equal number of days in the pollen season. Table 1 shows, as an example, the intervals which resulted in one study (15). This trial was held during a year when grass pollen counts were very low and thus none of the four intervals represents a very high level of pollen challenge to the patients. The "mean pollen count" is defined as the average of all the daily pollen counts which fall within the range of one quartile.

Table 1

Pollen count intervals in grains/m³ air/day

Quartile	Range of daily pollen counts	Mean pollen count for quartile
1	0–3	2
2	4–8	6
3	9–20	16
4	21–61	34

Symptom scores are then analysed for each pollen interval separately. Application of standard statistical techniques, for example, Student's *t* test, analysis of variance or covariance, or Mann-Whitney U test (1, 2, 14), will give mean (or median) values of symptom scores in each pollen interval for each treatment group. Using the mean pollen counts in Table 1, a response curve of symptom scores against pollen count can be drawn for each treatment group. This is analogous to a dose response curve. Fig. 5 shows two such response curves (15). In this example it is clear that not only is there no overall difference between active and placebo treatments, but an increase in pollen count from 2 to 34 grains/m³ air/day causes slightly more symptoms in the active treatment

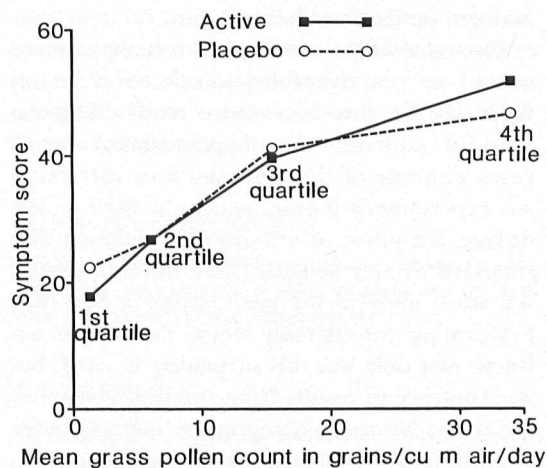


Fig. 5. Mean symptom scores for the four quartiles of the grass pollen season.

group than in the placebo group. This suggests that the patients receiving active treatment are slightly more sensitive to grass pollen than patients receiving placebo. This provides further evidence against the effectiveness of the active treatment. However, the opposite result might have occurred: the effectiveness of the active treatment in reducing symptoms might have been masked by a high level of symptoms which were not caused by grass pollen. The pollen count interval method of analysis would have shown this clearly.

DISCUSSION

Both of the methods described can be applied to either symptom or medication scores obtained from patient diaries, although results based on medications are more difficult to interpret.

The Normalised Regression method requires powerful computing facilities but it is an automatic method requiring no statistical control or judgement while it is being performed. This can be both an advantage and a disadvantage. The method is particularly useful when pollen counts are obtained from traps in more than one area, as occurs in multicentre trials. These are often necessary when two active treatments are being compared, because one centre is often unable to recruit the large number of patients necessary to detect relatively small differences between similar treatments.

The value of the method was recently demonstrated in a study of hyposensitisation in hay fever (7). In this four-centre study 58 grass-sensitive patients were hyposensitised for 3 years with one of the following four extracts: – An experimental 2-grass extract at high or low dosage, a 2-grass, or a 5-grass formulation of a standard extract. Results from the third year of the study showed the new extract at *high* dose performing considerably worse than other extracts. Not only was this surprising in itself, but was contrary to results from the first 2 years of the study. Normalised regression methods were applied, and the anomalous difference between treatment groups was found to be present in the analysis of the intercepts, but not the slopes

from the regression. In other words, the extra symptoms suffered by patients receiving the experimental extract at high dose were unrelated to the immediate effects of grass pollen. It was suggested that the extra symptoms were due to patients who were also birch sensitive entering the grass pollen season with hyper-reactive respiratory mucosae.

The inverse normal scores transformation (2) can be applied to results from all the years of a study grouped together. The slopes can then be compared between different years. In the above trial it was thus demonstrated that the third year of treatment resulted in a significant reduction in grass sensitivity (7).

The Pollen Interval method is much simpler in terms of computation and produces results which are easier to present than those from the Normalised Regression method. For these reasons it is probably to be preferred in a trial involving only one pollen count. An additional advantage is that Pollen Intervals can be useful when replacing missing data from a patient diary. Each missing value can be replaced by the mean of the values recorded by the patient on all other days in the same Pollen Interval.

When the trial design requires a number of pollen counts (for different centres and years of the trial) the choice of the pollen intervals can be difficult, particularly if one centre or year has a very high or low count compared with other centres or years. However, if care is taken in the design of the intervals and the interpretation of results, considerable insight can be gained into the relationship between symptoms and pollen counts.

In order to achieve the best results with either technique the pollen count must be of high quality, and must reflect accurately the pollen levels to which patients are actually exposed. Thus, diary data from any day on which a patient is not in the area of the appropriate pollen trap should be excluded from the analysis. Of even greater importance is the need to exclude all symptoms caused by exposure to other allergens, e.g. animal dander and caused by other diseases such as acute upper respiratory infections. We are aware that pollen exposure causing the allergic symptoms also de-

depends on whether the patient is inside or outside. This could be recorded on diary cards, but this information would be very difficult to compare from day to day. Outside, does that mean in the middle of a grass field (for a grass allergic patient) or in the middle of town? Inside, does that mean in a house with all windows shut or open? Because of these considerations we do not recommend that the patients should record being inside or outside during the day.

Neither of the two methods described represent a complete answer to all the problems in relating symptoms from patient diaries to pollen counts. Probably the main difficulty in both cases is the interpretation of symptoms which are not due to the immediate effects of the pollen count. These could be caused by other allergens, or be due to delayed effects of the pollen under study, or may be non-allergic symptoms caused by irritants, for instance smog and smoke. In a recent study (15) in grass sensitive patients, patients completed daily diaries for 3 months before the grass pollen season. Apart from those patients who were also tree sensitive, some other patients recorded fairly constant levels of nasal symptoms throughout these early months. These symptoms were diagnosed as being non-allergic in nature.

To get comparable symptom scores in clinical trials the anti-allergic medication must be standardized. Hay fever patients, for instance, should only be allowed to use one kind of antihistamine tablets or antihistamine eye and nose drops and, finally, prednisone tablets if allergic symptoms cannot be controlled by antihistamine alone. The daily quantity of medicine taken should be recorded on diary cards. Thus it is possible to calculate a reliable arbitrary medication score. If each patient only records the total number of times the medication is used, and not the total quantity of each medication used, misleading scores can result. Of even greater importance is that such a medication procedure will invalidate symptom score results, making it impossible to compare the individual scores accurately, since a higher medication in one of two comparable groups will cause fewer symptoms in the group receiving more antiallergic medicine. When performing a clinical placebo con-

trolled trial to show the efficacy of a new immunotherapy in, for instance, birch pollen hay fever, the patients in the placebo and the active treated group should be comparable as regards the severity of the allergic disease. If the groups are not comparable a new treatment might happen to have some effect, but it cannot be shown by comparing the symptom scores as shown in Fig. 5. In birch pollen allergic patients Viander et al. (16) have shown that the total seasonal symptom score is significantly correlated to the nasal sensitivity measured by nasal provocation test and the RAST result. As mentioned above, persons in such clinical trials should not be sensitive to allergens other than those under study in the observation period, but if so it is desirable that the groups are balanced concerning such allergies.

A method for evaluating therapy for hay fever was described recently (4). A crossover trial was carried out in 61 patients, comparing four treatments: Topical corticosteroids, topical cromoglycates, placebo tablets, and a new combination tablet. Three measurements were used to compare these treatments: a) post-treatment symptom questionnaires, b) daily analogue line scores of symptoms, and c) nasal FEV₁. Daily analogue line diary scores with an analysis which took into account variation in pollen count proved to be the most sensitive of the three measurements.

In conclusion, both the Normalized Regression method and the Pollen Interval method, when applied to symptom or medication scores in trials of hyposensitisation in hay fever, provide a better understanding of the effectiveness of the treatments involved.

ACKNOWLEDGEMENTS

We wish to thank Barry Blakesley and Larry Dineen for their work in the development of these methods, and Dr. Kirk Watson for all his help and encouragement.

REFERENCES

1. Armitage, P.: Statistical methods in medical research. Blackwell, Oxford, 1977.
2. Bradley, J. V.: Distribution-free statistical tests. Prentice-Hall, New Jersey, 1968.

3. Bruno, G.: Il (conjuvac) nella immunoterapia specifica della sindromi allergiche da Graminacee. *Folia Allergol. Immunol. Clin.* 29, 189–193, 1982.
4. D'Souza, M. F., Emanuel, M. B., Cregg, J., Charlton, J. & Goldschmidt, J.: A method for evaluating therapy for hay fever. *Clin. Allergy* 13, 329–335, 1983.
5. Frostad, A. B., Grimmer, Ø., Sandvik, L. & Aas, K.: Comparing a purified (refined) allergen preparation and a crude aqueous extract from timothy pollen. *Allergy* 35, 81–95, 1980.
6. Haahtela, T., Heiskala, M. & Suoniemi, I.: Allergic disorders and immediate skin test reactivity in Finnish adolescents. *Allergy* 35, 433–441, 1980.
7. Haahtela, T., Hagelund, C. H., Munck, E., Wihl, J.-Å. & Vilkkä, V.: Hyposensitisation. Comparing a purified alum-absorbed allergen preparation and a pyridine-extracted alum precipitated allergen extract (Allpyral®) from timothy and cocksfoot pollen. Read at XIV Nordic Allergy Congress, Gothenburg, 1981.
8. Horak, F. & Jäger, S.: Welchen fortschritt bringen modern konzipierte allergenlösungen? Ergebnisse einer zweijährigen multizentrenstudie mit Conjuvac. *Allergologie* 5, 290–298, 1982.
9. Lowell, F. C. & Franklin, W.: A double blind study of the effectiveness and specificity of injection therapy in ragweed hay fever. *N. Engl. J. Med.* 273, 675–679, 1965.
10. Norman, P. S.: A rational approach to desensitisation. *J. Allergy* 44, 129–145, 1969.
11. Norman, P. S., Lichtenstein, L. M., Kagey-Sobotka, A. & March, D. G.: Controlled evaluation of allergoid in the immunotherapy of ragweed hay fever. *J. Allergy Clin. Immunol.* 70, 248–260, 1982.
12. Østerballe, O.: Immunotherapy in hay fever with two major allergens 19,25 and partly purified extract of timothy grass pollen. *Allergy* 35, 473–479, 1980.
13. Pegelow, K.-O., Broman, P., Heilborn, H. & Sundin, B.: Immunoterapi med Allpyral och Conjuvac gräspollen-extrakt. Ann. Meeting of Swedish Medical Society, 1981.
14. Siegel, S.: Non-parametric statistics for the behavioural sciences. McGraw-Hill, New York, Toronto, London, 1956.
15. Taudorf, E., Laursen, L., Djurup, R., Kappelgaard, E., Pedersen, C. T., Søbørg, M., Wilkinson, P. & Wecke, B.: Oral administration of grass pollen to hay fever patients. An efficacy study in oral hyposensitization. (Submitted to *Allergy*).
16. Viander, M. R. & Koivikko, A.: The seasonal symptoms of hyposensitized and untreated hay fever patients in relation to birch pollen counts: correlations with nasal sensitivity, prick tests and RAST. *Clin. Allergy* 8, 387–396, 1978.

Address:

Ebbe Taudorf

Medical Dept. TA

Allergy Unit 7511

Rigshospitalet, Tagensvej 20

DK-2200 Copenhagen N

Denmark

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.