Europ. J. Cancer, Vol. 15, pp. 1–5

0014-2964/79/0001-0001/$02.00/0

*Perspectives in Cancer Research*

# The Use and Limitations of a Stratified Randomization in Surgical Adjuvant Studies of Cancer*

WILLEM K. AMERY† and JEAN DONY‡
*Janssen Pharmaceutica, N.V., B-2350 Beerse, Belgium*

**Abstract**—*Practical and theoretical limitations to the use of stratification in randomized clinical cancer studies are discussed in the light of the authors' practical experience in designing a multicenter adjuvant study with levamisole in resectable lung cancer. It is advocated that stratification be limited to only one or a very few variables and that an effort be made to stratify by hidden variables whenever this is feasible. In the latter context, more attention could be paid to a stratification by the co-operating centers, even if this carries the risk of an unequal distribution of patients regarding known prognostic variables as is the case with simple randomization itself. If such maldistributions do occur, a stratification of the analysis of the results will minimize the pitfalls potentially involved in the interpretation of the findings.*

## INTRODUCTION

IT APPEARS very fashionable nowadays in large-scale multicenter cancer studies to have the randomization of the patients stratified along certain prognostic variables. The main goal of this procedure is to increase the comparability of these variables in the treatment groups. Nothing is wrong with fashion, even in science, as long as it is the result of sound considerations. Is this the case with the stratification in cancer trials? And, if so, are there any limitations? These and related questions are considered in this paper, using our own experience in designing a protocol for a levamisole study in lung cancer as an illustration of some practical implications in stratifying a surgical adjuvant study.

## LIMITATIONS IN THE USE OF THE STRATIFICATION PROCEDURE

A wide variety of variables have been proved or are suspected of having a prognostic significance in cancer patients. These variables may be classified as follows:

1. Variables related to the tumor. These include growth characteristics (mitotic index; growth fraction; doubling time, etc.), the extent of the tumor (size of primary; presence, number and site of secondaries; the tumor "stage" encompassing most of the information provided by the two heretofore-mentioned variables; serum levels of tumor-associated antigens; serum $\beta_2$-microglobulin; etc.), the histological appearance of the tumor (histological classification; blood vessel invasion, etc.), and various others (such as: presence of hormone receptors and history of the disease, i.e. duration and type of symptoms).

2. Variables related to the host. Here we have the patients' general characteristics (age, sex, weight, etc.), variables concerned with the host's defence [(a) histological findings such as stroma reaction and sinus histiocytosis; (b) data on tumor specific immunity such as cellular cytotoxicity, cytotoxic antibodies and blocking antibodies; (c) direct measurements of immunity such as enumeration and functional status of T-, B- and null lymphocytes, monocyte-macrophage functions, skin tests, etc.; (d) measurements indirectly indicative of a disordered immunity, such as the presence of circulating immune complexes, of the rheumatoid factor and/or of acute phase reactants; (e) serum lysozyme levels; etc.], the general

*Presented in part at the "International Meeting on Comparative Therapeutic Trials" (Paris, 9–10 January 1978).

†Chairman of and ‡Consultant on Statistics to the Study Group for Bronchogenic Carcinoma, Postal address: Janssen Pharmaceutica N.V., B-2350 Beerse, Belgium.

and socio-economic status of the patient, and, possibly, hereditary factors.

3. Exogenous variables. Though most of these are often ignored, there is some indication that they may be influential: surgery and the treating center (anesthetics and other drugs used during anesthesia, the stress of surgery itself, the attitude of the surgeon and the nursing personnel), quantitative and qualitative consumption of medicines (such as antibiotics, anticoagulants, antiphlogistic drugs) and of certain food constituents, the presence of intercurrent infections, both localized (e.g. post-operative empyema) and generalized ones (e.g., influenza), and, perhaps, chronobiological and climatological factors.

The above list, although impressive, should not lead to the conclusion that stratification for prognostic variables serves an unattainable goal. It is, nevertheless, helpful in understanding the limitations of stratification in cancer studies. What are these limitations?

## 1. *Practical limitations*

These can be discussed in three broad categories:

(*a*) *The number of variables which one can use for stratification purposes.* This is limited because the eventual number of strata (i.e., subgroups for separate randomization) increases exponentially as the number of variables increases linearly.

Just to give one simple example: if one intends to use only five variables, which each split up the population into two equal groups, for stratification of the randomization one needs more than 300 patients to obtain strata of ten patients each. Obviously, as there is a limit to every population selected for a study, there will also be a limit to the number of variables used for stratification if one intends to end up with strata which are still large enough to keep the randomization performed within each stratum sensible.

(*b*) *Some variables are difficult to assess in an exact manner.* This is particularly true for variables in which subjective appreciation is involved. Examples of such variables are: disease history, stroma reaction, sinus histiocytosis, general health status, several variables related to the treating center and several others.

Such variables are, for that reason, not very suitable for subdividing the patients into two or more groups, because such a subdivision would be under suspicion of being based upon unreliable grounds.

(*c*) *Some variables are unknown.* Obviously unknown variables cannot be used for stratification. This does not mean, however, that we should forget about them entirely.

Variables may be unknown because they are hidden. Nobody will deny that such variables affect the prognosis of cancer patients, but as long as they remain unidentified they can hardly be used for stratification.

Other variables are only temporarily unknown. Here we find all relevant information that is obtained after randomization. Examples are the occurrence of localized or generalized infections, the use of certain drugs, and, in the particular case where the adjuvant treatment is started preoperatively, any information (including most pathology data!) which is collected at the time of surgery or thereafter.

## 2. *Theoretical considerations*

Any method other than randomization to assign patients to two or more treatment groups will be reliable only if all factors affecting prognosis are known, so that one can make sure that the treatment groups are identical in all important variables except the treatment under study [1]. It is clear that this requirement is not fulfilled in most clinical studies and certainly not in cancer. Therefore, there is no valid alternative available to randomization in cancer studies to-day, all the more so since it is precisely when there are hidden variables which may be influential that randomization is important [2]. It follows that stratification does not aim at making the randomization senseless or superfluous, but that it should serve as a sensible means of manipulating the randomization procedure.

Randomization is reasonably reliable only when the number of patients used is large enough. The larger the number, the smaller the chances are to end up with an unequal distribution of patients. Stratification, however, results in the creation of subgroups (the "strata") for randomization, and, as discussed above, increasing the number of stratification variables readily results in strata which risk being too small to ensure reliability of the randomization procedure itself. These considerations argue against the use of an elaborate stratification and clearly suggest that using one or only very few variables is superior to using several ones; this is especially so if the random allocations within the strata

are balanced, so that numbers on different treatments within any one stratum remain nearly equal.

## FROM THEORY TO PRACTICE

One is tempted to conclude from the above, that stratification, being a means but not a goal in itself, should be restricted to one or only a few available variables which ensure the formation of sufficiently large strata. Moreover, if possible, any measure which may be helpful in avoiding maldistribution of hidden variables should be taken.

We tried to translate these principles into practice when we were designing a protocol for a multicenter placebo-controlled adjuvant study with levamisole a few years ago. The factors which we had to take into account were the following:

1. Prognostic variables as described in the literature.

2. The number of patients available for the study within a reasonable time was estimated to be around 200.

3. There were reasons to start this type of immunotherapy before the operation. As discussed at more length elsewhere [3, 4], these reasons were the following:

(i) An early start of immunotherapy seems warranted because immunity is thought to only be able to deal with a small number of cancer cells.

(ii) Levamisole restores the host defence mechanisms when these are deficient [5]. As surgery is associated with a temporary, reversible immune suppression, it seemed preferable to the investigators to start levamisole treatment before the operation in an attempt to prevent this type of immune suppression.

In view of these elements, we decided to restrict the stratification to one single variable. Whilst reviewing the list of prognostic variables in resectable lung cancer, we found that:

(a) Commonly used variables related to the extent of the tumor or its histological type would not preoperatively be available in a substantial number of patients. This type of variables could thus hardly be used.

(b) Other variables, such as age, which sometimes are used for stratification purposes, had only limited or no bearing at all on the prognosis of such patients.

(c) Several of the hidden variables could be related in some way to the treating center, since it is well known by people who are experienced in multicenter trials that differences do occur between the co-operating centers in such studies and that it appears extremely difficult to eliminate such differences.

It occurred, therefore, to us that stratifying the randomization by the (three) co-operating centers could provide the best solution to the problem as it resulted in strata which were reasonably large enough and it could be expected to increase the comparability of the two treatment groups regarding certain hidden variables. The obvious risk in this was that, as is the case with any unstratified randomized trial, the treatment groups could appear to show some extent of maldistribution regarding "known" prognostic variables. Therefore, we specified in the protocol that during the analysis of the data, a search be made for such potential maldistributions especially those concerned with the extent of the tumor, as this appeared the most decisive factor amongst the known prognostic variables. If any maldistribution would occur, we planned to make an analysis stratified along the variable concerned, in order to find out whether a potential difference between the two treatment groups was the consequence of that maldistribution. It has been argued that such a stratification of the analysis—in other words: a stratification after selection—is almost as efficient as an ordinary pre-randomization stratification [6–8].

The protocol has been described and discussed in detail elsewhere [3, 4] and therefore will not be repeated. The study is approaching its end and we felt that time had come to confine our experience, regarding the problem of stratification, to paper. When, in May 1977, the latest interim evaluation of the results was performed [9], all 211 patients had been in the study for more than 1 yr. Levamisole appeared superior to the placebo regarding the end-points of the study, i.e., the rate of recurrences and of carcinomatous deaths, and this trend was detected in the three co-operating centers. However, it appeared that the randomization could have favored the levamisole group because of a somewhat higher number of placebo patients with more advanced cancers ($0.05 < P < 0.10$). We had been using two measures to assess the tumor extent, i.e., the largest diameter of the primary lesion after its fixation by the pathologist and a grouping system for the regional extent of the tumor. The details of the latter grouping system will not be repeated here: it

will suffice to say that Category 1a, the prognostically most favorable group, encompassed all those patients in whom the cancer appeared to be limited to its primary site but in whom no evidence of blood vessel invasion could be detected. We, therefore, stratified our analysis along the two variables of tumor extent. The result of this is given in Fig. 1. It

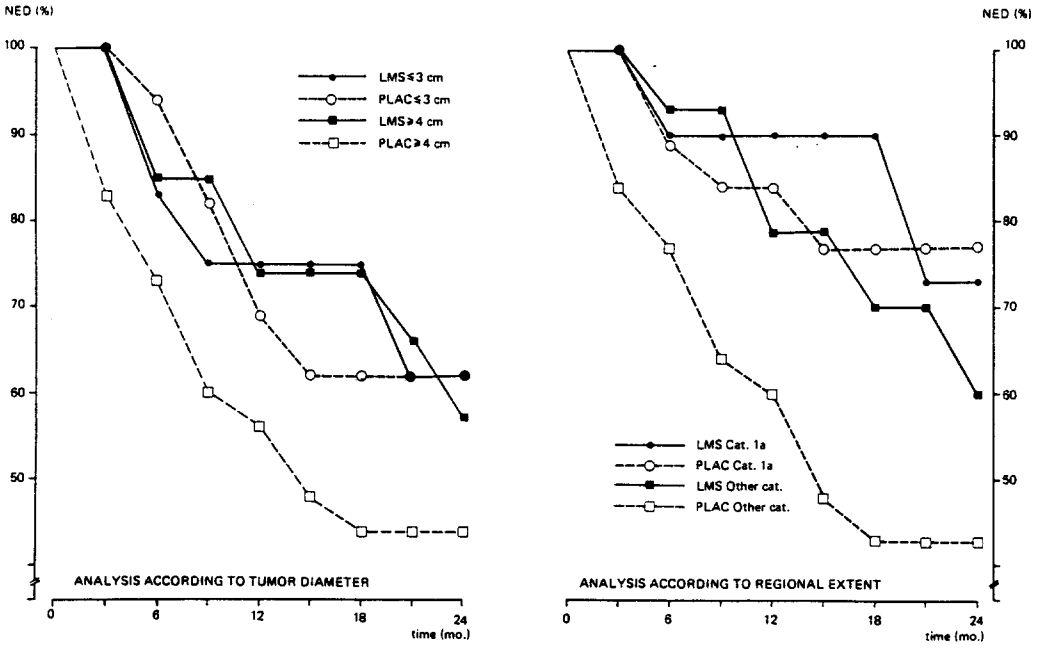


*Fig. 1. Remission duration (actuarial analysis) as related to tumor extent in adequately treated patients.*

appears from this analysis that the difference between levamisole and the placebo treatment was actually most pronounced in those patients who had the more advanced cancers, whereas such a difference was far from marked in the others. Such a trend (clearer superiority if the pre-treatment tumor load is heavier) has been found in the meantime in other controlled studies, as well [10]. It would seem from the above, therefore, that the difference between levamisole and the placebo can hardly be explained by a maldistribution of the patients regarding tumor extent at the start.

## CONCLUSION

The moral from our experience and the tentative conclusion from this presentation could be that one should limit oneself to only a (very) few variables if stratification is planned and that, whenever possible one should aim to stratify along hidden variables. As certain of these variables may be grouped in the various co-operating centres, we are tempted to advocate that, in future cancer studies, more attention be paid to a stratification by these centers. This still carries the risk of resulting in some degree of maldistribution of other known variables, but such maldistributions can be (and: should be) checked afterwards, which is not feasible regarding the hidden factors. If such a maldistribution is found, the analysis ought to be stratified along the variable concerned. This offers the advantage of limiting the difficulties in the eventual interpretation of the results.

**Acknowledgements**—The authors are indebted to Dr. Paul Janssen and Mr. Ludo Gypen, for reviewing the manuscript and to Mr. Hugo Vanhove for brushing-up the language of the text.

## REFERENCES

1. D. P. Byar, R. M. Simon, W. T. Friedewald, J. J. Schlesselman, D. L. de Mets, J. H. Ellenberg, M. H. Gail and J. H. Ware, Randomized clinical trials. Perspectives on some recent ideas. *New Engl. J. Med.* **295,** 74 (1976).
2. P. Meier, Statistics and medical experimentation. *Biometrics* **31,** 511 (1975).

3. STUDY GROUP FOR BRONCHOGENIC CARCINOMA, Immunopotentiation with levamisole in resectable bronchogenic carcinoma: a double-blind controlled trial. *Brit. med. J.* **3,** 461 (1975).
4. W. K. AMERY, Double-blind levamisole trial in resectable lung cancer. *Ann. N.Y. Acad. Sci.* **277,** 260 (1976).
5. J. SYMOENS and M. ROSENTHAL, Levamisole in the modulation of the immune response: the current experimental and clinical state. *J. Reticuloendothel. Syst.* **21,** 175 (1977).
6. P. ARMITAGE and E. A. GEHAN, Statistical methods for the identification and use of prognostic factors. *Int. J. Cancer.* **13,** 16 (1974).
7. M. G. KENDALL and W. R. BUCKLAND, *A Dictionary of Statistical Terms.* Oliver and Boyd, Edinburgh (1957).
8. R. PETO, M. C. PIKE, P. ARMITAGE, N. E. BRESLOW, D. R. COX, S. V. HOWARD, N. MANTEL, K. MCPHERSON, J. PETO and P. G. SMITH, Design and analysis of randomized clinical trials requiring prolonged observation of each patient. *Brit. J. Cancer.* **34,** 585 (1976).
9. W. K. AMERY, Levamisole in resectable lung cancer. *Symposium on Immunotherapy of Malignant Diseases*, Vienna, 9–10 November (1977).
10. W. K. AMERY, Overview of other controlled clinical data: an attempt at defining the future position of levamisole in cancer therapy. *Symposium on Immunotherapy of Malignant Diseases.* Vienna, 9–10 November (1977).