CLINICAL TRIAL OF THE THERAPEUTIC VALUE OF CERTAIN ANTIRHEUMATIC DRUGS

BY

P. VAN DER MEER AND S. A. DEN OUDSTEN

Consultatiebureaus voor Rheumatische Ziekten of the Rotterdamse Stichting voor Rheumabestrijding, Rotterdam, Netherlands

In addition to the so-called basic therapy employed in the treatment of rheumatoid arthritis, a number of drugs are in general use. Many years of experience have shown the value of aspirin, aminopyrin (pyramidon), phenylbutazone (butazolidin), and gold. This evaluation is based primarily on estimates of their antirheumatic and analgesic effects, as well as on considerations of the risk of toxicity.

It is not proposed to discuss the possibility of a distinct differentiation between the analgesic and antirheumatic effects of these drugs on the one hand and the antirheumatic and antiphlogistic properties on the other. This subject has been fully treated by Winder (1959).

A better understanding of the course of rheumatoid arthritis treated exclusively with basic therapy has established that, because of the variable and unpredictable course of the disease, any evaluation of antirheumatic drugs must satisfy a number of demands. Such an evaluation can be carried out most successfully, by methods which have been worked out largely by British investigators, in so-called "double-blind" trials, by which neither the patient nor the investigator knows which drug is being administered.

A double-blind test with salicylamide and a placebo has been reported by Batterman and Grossman (1955), but no similar study of the analgesic properties of aminopyrin has come to the authors' attention. Five blind investigations of phenylbutazone have been reported (Steinbrocker, Berkowitz, Carp, Ehrlich, and Elkind, 1951; Ansell, Fearnley, Bywaters, and Meanock, 1953; Fjellström, Goldberg, Lindgren, and Nilsson, 1957; Smyth and Clark, 1957; Vaughn, Howell, and Kiern, 1959). The method used by Hajnal, Sharp, and Popert (1959), who tested aspirin, prednisone, and phenylbutazone, is not comparable with the usual double-blind type of trial.

Our object was not only to assess the clinical value of aspirin, aminopyrin, and phenylbutazone in relation to a placebo, but at the same time to compare these drugs with each other as regards therapeutic activity and toxicity.

Method

Material.—191 patients from the Clinical Centres for Rheumatic Diseases N and S (Rotterdam) with "definite rheumatoid arthritis" (A.R.A. criteria: Ropes, 1959) received, in addition to the basic treatment, a 6-week course of drugs in the form of powders. The choice of the first powder to be administered was determined by a random list, as was the order in which the drugs and the placebo were taken by the patient.

Disease Activity.—This was measured before and after the course of treatment according to four criteria:

1. Strength of grip (mm. Hg), measured by a standardized technique (average for both hands).

2. Pain on firm digital pressure over the affected joint (mostly the metacarpophalangeal), measured in four grades:
   - 0 = No pain
   - 1 = Slight pain
   - 2 = Severe pain (with wincing)
   - 3 = Pain (with wincing and withdrawal)

3. Erythrocyte sedimentation rate (mm./hr, Westergren).

4. Patient’s subjective opinion as to his condition at the end of the 6 weeks, using the following ratings:
   - 0 = Unchanged
   - 1 = Worse
   - 2 = Much worse
   - 3 = Somewhat improved
   - 4 = Greatly improved

Grouping the Patients.—The three drugs and the placebo were identified by the numbers I, II, III, IV, and they were handled under these designations throughout the experiment, the names not being disclosed even to the authors until the end of the trial.
The 191 patients were grouped according to the number of the first drug used. The clinical details as at the start of the experiment are shown in Table I.

**Dosage.**—The following dosages were used:
- Phenylbutazole, three daily doses of 200 mg.;
- Aminopyrin, 500 mg., three times daily;
- Aspirin, 1 g., three times daily (at the start of our trial it was not sufficiently known that the dose should have been at least 6 g. to be compared with 600 mg. phenylbutazole);
- Sach. lact., 1 g., three times daily.

The powders were all given the same bitter taste by the addition of a small amount of quinine.

2 weeks after the conclusion of one course of powders the next drug in the series was administered for a further 6 weeks. In this way it was possible to compare the effect of various drugs on various patients, and all also received the basic treatment throughout the investigation.

**Toxicity.**—When toxic symptoms appeared, an attempt was made to carry on with a smaller dosage, but the administration was discontinued if toxic signs persisted, and a few patients defaulted.

In connexion with the danger of intoxication, all patients were checked weekly during the consulting hour as to their subjective complaints. Weekly procedures included a leucocyte count, and the urine was checked for albumen, urobilin and sediment.

**Results**

Table II shows the percentage of patients who completed courses of each drug. The chances of completing the course decreased in the sequence I-II-III-IV; in other words, Drug IV seems to have been more toxic than Drug I; the differences in frequency of completion, however, are not significant.

Table III (opposite) shows the reasons for discontinuing courses. The most important toxic reactions were the following:

1. **Decrease in the Number of Leucocytes** (lower than 4,000/c.mm.).
   - Drug II—0
   - Drugs I and III—5 each
   - Drug IV—6.

2. **Oedema**.
   - Drug IV—6.

3. **Stomach Complaints**.
   - Drug II—6
   - Drug III—13
   - Drug IV—8
   - Drug I—5

The numbers of separate complaints are too small for the differences to be significant.
TRIAL OF ANTIRHEUMATIC DRUGS

TABLE III

REASONS FOR DISCONTINUING COURSES

<table>
<thead>
<tr>
<th>Drug No.</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Itching</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Decrease in Number of Leucocytes (below 4,000)</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Stomach Complaints</td>
<td>5</td>
<td>16</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Throat Complaints</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Gastro-Intestinal Bleeding</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dizziness</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Exanthema</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Abdominal Complaints</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Ringing in the Ears</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Thrombophlebitis</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Icterus</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Oedema</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Headache</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total No.</td>
<td>16</td>
<td>23</td>
<td>32</td>
<td>35</td>
</tr>
<tr>
<td>Per cent.</td>
<td>17</td>
<td>24</td>
<td>27</td>
<td>35</td>
</tr>
</tbody>
</table>

Patients Defaulted | 13 | 9 | 14 | 8 |

Total Discontinued | 29 | 33 | 46 | 43 |
| Per cent. | 32 | 34 | 39 | 43 |

It is also possible, since so many of the complaints were subjective in nature, that they came from patients who invariably complained; and indeed the percentage of persons who discontinued the experiment of their own accord (i.e. not on the advice of their physician) decreased with the number of courses taken: 49 per cent. having taken one course, 45 per cent. two courses, 32 per cent. three courses, and 28 per cent. four courses. This suggests the operation of a certain psychological factor.

There was no indication that the frequency of discontinuation depended on the drug with which the series was begun.

Table IV gives the means of the results obtained by evaluating the courses of each drug separately. Table I shows that the four groups to which patients were allotted at random agreed well as to age, sex, duration of disease, functional capacity, erythrocyte sedimentation rate, grip, and pain on pressure. It must therefore be concluded that the difference in the results was caused by differences in the effects of the drugs.

As not all persons received a complete course with each of the four drugs it was deemed necessary to compare also the effect of the drugs in pairs, using the results of only those persons who were treated with both; these comparisons were made separately for grip, pressure pain, erythrocyte sedimentation rate, and subjective evaluation. These detailed comparisons were made as follows:

<table>
<thead>
<tr>
<th>Drugs</th>
<th>No. of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>I and II</td>
<td>34</td>
</tr>
<tr>
<td>I and III</td>
<td>28</td>
</tr>
<tr>
<td>I and IV</td>
<td>17</td>
</tr>
<tr>
<td>II and III</td>
<td>36</td>
</tr>
<tr>
<td>II and IV</td>
<td>19</td>
</tr>
<tr>
<td>III and IV</td>
<td>27</td>
</tr>
</tbody>
</table>

In Table V (overleaf) the result of each comparison is given by an arrow. The points of the arrows indicate the best results. A double arrow is used when the difference is significant at the 5 per cent. level. Below the diagram the drugs are shown in order of effectiveness.

Table VI (overleaf) compares both sequences, namely that obtained by mutual comparison (Table V) and that obtained by evaluating the courses of each drug separately (collective average, see Table IV).

Grip.—The sequence in both Tables is I, III, II, IV.

Pain.—There is agreement for I and III and disagreement for II and IV.

Erythrocyte Sedimentation Rate.—There is no agreement at all.

Subjective Improvement.—Complete agreement is seen in the sequence I, II, III, and IV, in the sense that IV is significantly higher than the rest.

Discussion

There is therefore some disagreement between the sequence found by comparing the drugs in pairs and that found by comparing the means for all observations with each other. Nevertheless,
TABLE V
MUTUALLY COMPARED THERAPEUTIC ACTIVITY

<table>
<thead>
<tr>
<th>Grip</th>
<th>Pressure Pain</th>
<th>Sedimentation</th>
<th>Subjective Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
<td><img src="image3" alt="Diagram" /></td>
<td><img src="image4" alt="Diagram" /></td>
</tr>
</tbody>
</table>

Sequence According to Diagram:
I-III-II-IV
I-III-IV-II
I-II-IV-III
I-II-III-IV

The result of each comparison is given by an arrow; the points of the arrows indicate the best results. A double arrow is used when the difference is significant at the 5 per cent. level.

TABLE VI
MUTUALLY COMPARED THERAPEUTIC ACTIVITY

<table>
<thead>
<tr>
<th>Grip</th>
<th>Pressure Pain</th>
<th>Sedimentation</th>
<th>Subjective Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5" alt="Diagram" /></td>
<td><img src="image6" alt="Diagram" /></td>
<td><img src="image7" alt="Diagram" /></td>
<td><img src="image8" alt="Diagram" /></td>
</tr>
</tbody>
</table>

Sequence According to Diagram:
I-III-II-IV
I-III-IV-II
I-II-IV-III
I-II-III-IV

Sequence According to Collective Average:
I-III-II-IV†
I-III*-II†-IV†
IV-I-III-II*
I†-II†-III†-IV†

* Significant at the 5 per cent. level.
† Significant at the 1 per cent. level.

In general, Drug I gives the least effect and Drug IV the most (excluding the erythrocyte sedimentation rate).

At this point the reader will be curious, as indeed were the authors, at the conclusion of the experiment and the statistical analysis, to know which number stood for which drug. The code was as follows:

- I = Placebo
- II = Aspirin
- III = Aminopyrin (Pyramidon)
- IV = Phenylbutazone.
Conclusion

The placebo as well as the drugs had, as was to be expected, a favourable effect on the symptoms of rheumatoid arthritis. This must be ascribed to the happily ever-present tendency to spontaneous recovery from rheumatoid arthritis, and also to the effect of suggestion which is inherent in all placebos.

Aspirin, aminopyrin, and phenylbutazone, however, had an appreciably better effect than the placebo, and of these phenylbutazone gave the best results. The authors wish to confine their conclusions from this experiment to the statement that, in this trial (bearing in mind the inadequacy and limitations of the quantitative criteria), phenylbutazone proved to be a better analgesic drug in cases of rheumatoid arthritis, than aminopyrin or aspirin in the dosages used.

The number of discontinued courses indicates that the administration of a placebo may be followed by the signs and symptoms of toxicity and intolerance (Table III). The clear-cut decrease in the number of leucocytes in patients taking the placebo must certainly have been due to a spontaneous fluctuation in the leucocyte count, but it led the authors, with exaggerated caution, to terminate the course.

The fact that toxicity increased in the sequence placebo, aspirin, aminopyrin, and phenylbutazone points to the conclusion that phenylbutazone is most likely to cause toxic reactions, although the differences were not large enough in these series to be significant.

Summary

In a “double-blind” trial of drugs in the treatment of rheumatoid arthritis, 191 patients with definite rheumatoid arthritis (American Rheumatism Association criteria) were given in four successive courses of 6 weeks: three daily doses of 1 g. aspirin, three daily doses of 0·5 g. aminopyrin, three daily doses of 200 mg. phenylbutazone, and three daily doses of a placebo.

All the powders looked the same and had the same taste. The order in which the drugs and the placebo were taken by the patient was determined by chance. Between each period of 6 weeks there was an interval of 2 weeks.

All the patients were examined at the beginning and end of each period by one of the authors and the improvement or deterioration was noted according to the following criteria:

1. Hand grip (mm. Hg);
2. Tenderness on pressure (metacarpophalangeal joints);
3. Erythrocyte sedimentation rate (Westergren);
4. Subjective judgment of the therapeutic result.

The results of this trial suggest that, in the dosages used, phenylbutazone is the best analgesic drug in rheumatoid arthritis, followed by aminopyrin and aspirin in that order, and the authors desire to limit their conclusions to this statement. Toxic reactions were least with the placebo, and increased with aspirin, aminopyrin, and phenylbutazone in that order, but the differences are not large enough to be significant in this series of patients.

The authors acknowledge with gratitude the co-operation of Dr. E. F. Drion and Mr. Th. J. D. Erlee of the Statistical Department of Toegepost Natuurkundig Onderzoek, The Hague, who did all the statistical work essential for this study, and gave much valuable criticism and advice.

REFERENCES


Essai en clinique de la valeur thérapeutique de certains médicaments antirhumatismaux

RÉSUMÉ

En un essai à “double feinte” (double-blind) de l’efficacité de certains médicaments dans le traitement de l’arthrite rhumatismale, 191 malades atteints de cette maladie (selon les critères de l’American Rheumatism Association) recurent, en quatre cures successives de 6 semaines, trois fois par jour, une des substances suivantes: 1 g. d’aspirine, 0·5 g. d’aminopyrine, 200 mg. de phenylbutazone ou une poudre inerte.

L’aspect et la saveur de toutes les poudres était la même. L’ordre d’administration des médecaments et de la poudre inerte à chaque malade fut déterminé par hasard. Chaque période de 6 semaines fut séparée par 2 semaines sans traitement.

Tous les malades furent examinés au début et à la fin de chaque période de traitement par un des auteurs et l’amélioration ou la détérioration fut déterminée selon les critères suivants:

1. La force de la poigne (mm. de Hg)
2. Douleur à la pression (articulations métacarpophalangiennes)
3. Vitesse de sédimentation globulaire (Westergren)

Les résultats de cet essai indiquent que le phenylbutazone est le meilleur des analgésiques dans l’arthrite rhumatismale, suivi par l’aminopyrine et l’aspirine dans cet ordre, et les auteurs désirent limiter leurs conclusions à cela. En ce qui concerne les réactions toxiques, elles furent minimes avec la poudre inerte et augmentèrent avec l’aspirine, l’aminopyrine et la phenylbutazone dans cet ordre, bien que les différences ne soient pas suffisantes pour être significatives dans cette série de cas.
Prueba clínica sobre el valor terapéutico de ciertos medicamentos antirreumáticos

Sumario

En una prueba a “doble disfraz” (double-blind) de la efectividad de ciertos medicamentos en el tratamiento de la artritis reumatoide, 191 enfermos padeciendo de tal enfermedad (de acuerdo con los criterios de la American Rheumatism Association) recibieron, en cuatro periodos sucesivos de 6 semanas: 3 dosis diarias de 1 g. de aspirina, 3 dosis diarias de 0,5 g. de aminopirina, 3 dosis diarias de 200 mg. de fenilbutazona y 3 dosis diarias de un polvo inerte (“placebo”).

Todos los medicamentos fueron administrados en forma de polvo y todos ellos eran iguales a la vista y tenían el mismo sabor. El orden en que los medicamentos y el “placebo” fueron administrados al paciente se determinó al azar. Entre cada periodo de 6 semanas de tratamiento se intercalaron 2 semanas de descanso.

Todos los enfermos fueron examinados al principio y al final de cada periodo de tratamiento por uno de los autores y la agravación o la mejora se determinó con arreglo a los siguientes criterios:
1. Fuerza al asir con la mano (mm. de Hg)
2. Dolor provocado a la presión (articulaciones metacarpofalángicas)
3. Velocidad de sedimentación eritrocitaria (Westergren)
4. Juicio subjetivo del resultado terapéutico.

Los resultados de esta prueba sugieren que la fenilbutazona es el mejor de los analgésicos en la artritis reumatoide, seguida por la aminopirina y la aspirina en este orden, deseando los autores limitar sus conclusiones a la expresada. En cuanto a las reacciones tóxicas fueron mínimas con el “placebo” y aumentaron con la aspirina, aminopirina y fenilbutazona en este orden, aunque las diferencias no fueron suficientes para ser significativas en esta serie de casos.