Clinical Evaluation of Some Phosphorus Bone-Imaging Agents: Concise Communication

Manno F. Nelson, L. Clifford McKee, and John R. Van Wazer
Vanderbilt University, Nashville, Tennessee

Four different Tc-99m labeled phosphorus-based bone-imaging agents were compared by a scan-evaluation technique in which the skeletal uptake was visually assessed at selected sites. A detailed statistical analysis of scans, by each of three independent evaluators, on a total of 140 different patients showed that two of the agents (based on pyro- and on trimetaphosphate) were superior. The statistical analysis showed good agreement between the investigators and consistency in their repeat evaluations.


The recent advent of Tc-99m labeled bone-imaging agents based on various phosphorus-containing compounds has initiated considerable discussion as to the efficacy of each agent used. Several articles have attempted to make comparisons; some evaluated the agents by following blood levels and determining uptake in selected bone sites, and others by inspection of scans (1—3). In several cases, comparison of two agents in the same select patient (1,3,4) was carried out. However, in each instance, the studies did not give sufficient data to permit statistically sound comparisons, or they failed to consider experimental factors that might possibly affect the results. We therefore designed a study that would be amenable to a thorough statistical analysis whereby comparison could be made between the bone-uptake properties of the three commercially available types of bone-imaging agents—the pyrophosphate, polyphosphate, and diphosphonate formulas—and those of the trimetaphosphate bone-seeker developed by us utilizing sodium (5). In this investigation, scans of hospital patients suspected of having bone lesions were evaluated with respect to the overall skeletal uptake as evidenced at selected sites, regardless of the presence or absence of bone abnormalities.

METHODS AND MATERIALS

The bone-imaging agents used in the studies were disodium etidronate,* tin pyrophosphate,† polyphosphate,‡ and our sodium trimetaphosphate formulation using 25 mg of phosphate and 0.5 mg stannous ion generated electrolytically. The pertechnetate Tc-99m used was obtained from Mo-99 generators and was alternately evacuated and nitrogen-purged until an aliquot added to an equal volume of acidified 0.1% sodium iodide solution gave no color with fresh starch solution. Each new generator was eluted at least once after receipt and before being eluted for use in the preparations. Each agent was prepared in compliance with the kit manufacturer's instructions and was used for all patients studied during a daily period. None of the patients had received a previous dose of any of the Tc-labeled bone-imaging agents. Before use, each radiopharmaceutical preparation was checked by thin-layer chromatography for pertechnetate content, and any preparation containing greater than 4% of unreacted pertechnetate was disqualified for this study. All administrations were done 45—70 min after preparation of the radiopharmaceutical, and all scans were run 3.5—5 hr after the injection. A new dual-head scanner fitted with 5-in. collimators was used in the studies. The patient's sternum was selected as the "set up" site for adjusting scanner controls.

Preparation of the various agents was completely randomized in respect to the day of the week; there was no sequence as to which agent would be used on which day. This was done to ensure that effects

* Received June 4, 1976; revision accepted Jan. 1, 1977.
For reprints contact: John R. Van Wazer, Vanderbilt University, Nashville, TN 37235.
such as radiolysis of the generator eluates or potential buildup of Tc-99m would be minimal. All scans collected on consecutive days over a period of several months were evaluated with respect to the scan quality at four specific locations, and overall, on each scan in the posterior view, as well as at four locations on each scan in the anterior view. The posterior checks consisted of (A) the ratio of backbone uptake to that of the lumbar–paravertebral void, (B) rib definition, (C) definition of the clavicle and the upper border of the scapula, and (D) definition of the iliac fossa, as well as (E) the overall lack of body background (soft-tissue and vascular-system content). Anterior locations evaluated consisted of (A) the contrast between femur and femoral-artery uptake, (B) definition of the femur, (C) definition of the sternum, and (B) rib definition. Each location was given either a high, medium, or low rating: +, 0, and —, respectively. Scans were rated blindly, with no chronologic succession in the scans evaluated and with the identity of the agent used unknown to the observer. Scans on 40–43 different patients were independently evaluated, by each of the three critics, for each of the following: the commercial diphosphonate kit, the pyrophosphate kit, and the trimetaphosphate formulation. Because of poor-quality scans, only 15 patients were similarly studied with the polyphosphate kit. In all, posterior and anterior scans of about 140 patients were independently rated by three professionals with respect to the aforementioned criteria. Furthermore, to determine the reproducibility of each of the three raters in his overall evaluation of each scan, a single set of randomly selected scans from 20 patients was rated blindly at two different times several weeks apart by each rater, with the ordering of the scans being different and randomly chosen for each rating. The resulting two sets of 20 overall evaluations were then intercompared for each rater.

One-way statistical analysis of variance without interaction—including the results of the three raters evaluating all four agents—was conducted using the method described by Freund and by Guenter (6,7). A one-way analysis of variance for three raters evaluating the three better agents was also conducted using Freund’s procedures (6). A two-way statistical analysis of variance (without interaction) was then carried out, so that the statistical variability of the four agents was determined along with the variability between the three evaluators (6).

RESULTS

Table 1 shows a typical set of one evaluator’s results on nineteen different patients, randomly se-

---

**Table 1. Blind Comparison of Bone-Visualizing Agents in Humans: An Evaluation of Eighteen Patients by One Rater**

<table>
<thead>
<tr>
<th>Patient</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>α</th>
<th>β</th>
<th>γ</th>
<th>δ</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>3542</td>
<td>1</td>
<td>+</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>—2</td>
</tr>
<tr>
<td>8949</td>
<td>2</td>
<td>—</td>
<td>—</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>—</td>
<td>—6</td>
</tr>
<tr>
<td>1457</td>
<td>3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+9</td>
</tr>
<tr>
<td>6578</td>
<td>4</td>
<td>+</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+4</td>
</tr>
<tr>
<td>8761</td>
<td>5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>—</td>
<td>0</td>
<td>—</td>
<td>+1</td>
</tr>
<tr>
<td>1044</td>
<td>6</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>+6</td>
</tr>
<tr>
<td>9713</td>
<td>7</td>
<td>+</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>+</td>
<td>—</td>
<td>0</td>
<td>—</td>
<td>+0</td>
</tr>
<tr>
<td>4442</td>
<td>8</td>
<td>+</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>+2</td>
</tr>
<tr>
<td>7391</td>
<td>9</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+5</td>
</tr>
<tr>
<td>5947</td>
<td>10</td>
<td>—</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>+7</td>
</tr>
<tr>
<td>4873</td>
<td>11</td>
<td>+</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>+2</td>
</tr>
<tr>
<td>4425</td>
<td>12</td>
<td>+</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>+3</td>
</tr>
<tr>
<td>2090</td>
<td>13</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+7</td>
</tr>
<tr>
<td>3633</td>
<td>14</td>
<td>+</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+1</td>
</tr>
<tr>
<td>6728</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>7721</td>
<td>16</td>
<td>+</td>
<td>—</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+4</td>
</tr>
<tr>
<td>8695</td>
<td>17</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+4</td>
</tr>
<tr>
<td>7759</td>
<td>18</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+3</td>
</tr>
</tbody>
</table>
lected with respect to the choice of bone-imaging agent used. This illustrates the standard rating sheet employed in this study. From the data obtained on all 140 patients by the three raters, the algebraic sums of the scan values for the four agents, averaged over all of the patients and the three raters, were: polyphosphate, -2.31; diphosphonate, +0.99; pyrophosphate, +2.21, and trimetaphosphate, +2.33.

Evaluation of the ability of all three raters to duplicate their own evaluations for scans on 20 patients was examined. Even though each investigator infrequently assigned a +1 in one reading and a -1 in the duplicate reading for a single evaluating location on a scan, the algebraic sum of the ratings at all nine locations per patient was usually the same for the pair of duplicate evaluations per patient per investigator. Furthermore, these duplicate values of the algebraic sum were found never to differ by more than one unit, although the maximum difference could have been eighteen units (i.e., +9 compared with -9). This means that the reproducibility by which each evaluator rated the given set of scans was very good.

The one-way statistical analysis of the results on all four agents, as obtained by the three evaluators, gave an F value of 31.89. A table of the F distribution for 3 and 8 degrees-of-freedom gave F_{.05} = 4.07 and F_{.01} = 7.59, respectively. Note that the 3 degrees-of-freedom derive from the four different bone-imaging agents tested, while the 8 degrees-of-freedom come from consideration of these four agents along with the variability between the three readers (see Ref. 6, p. 342). The large F value observed shows that between the four bone-imaging agents there is a difference that is statistically significant at the 0.01 level (i.e. the 99% critical value has been exceeded). A similar one-way analysis for the pyrophosphate, diphosphonate, and trimetaphosphate agents (with the polyphosphate omitted), and using the data from all three evaluators, gave an F value of 5.18. The table of F values for 2 and 6 degrees-of-freedom gave F_{.05} = 5.14 and F_{.01} = 10.9. These values show that the three agents are indistinguishable at the 0.01 level but distinguishable at the 0.05 level. From examination of the data, it is seen that the pyrophosphate and the trimetaphosphate are indistinguishable at the 0.05 level.

We next undertook a two-way analysis of the statistical variability of all four agents in comparison with the statistical variability between the three evaluators. An F value of 37.7 was obtained for the four agents compared to tabular F_{.05} and F_{.01} values of 4.76 and 9.78, respectively, for 3 and 6 degrees-of-freedom. In contrast, an F value of 1.73 was obtained for the scan readers compared to tabular F_{.05} and F_{.01} values of 5.14 and 10.9. As would be expected, the two-way analysis agrees with the one-way analysis of variance in that at least one of the four agents (specifically the polyphosphate) is clearly distinguishable. Furthermore, evaluations by the different raters are not significantly different even at the 0.05 level.

**DISCUSSION**

The statistical analyses that were carried out on the scans lead to the following conclusions:

1. The analyses of variance, taken together with the scan values, show that the polyphosphate is clearly different from, and inferior to, the diphosphonate, pyrophosphate, and trimetaphosphate formulations.

2. Likewise, the diphosphonate is statistically distinguishable as a less effective bone-imaging agent than the pyrophosphate or the trimetaphosphate.

3. The two-way analysis of variance demonstrates that the evaluations of each of the three scan readers are in statistical agreement with each other.

4. Blind duplicate evaluations of scans on 20 different patients showed that each scan reader could repeat his overall evaluation of a patient within 1 rating point even though the ratings could (and did) differ by a maximum of 18 points.

These results show the value of our rating technique in establishing an objective test for comparing bone-imaging agents. The ability of the individual scan readers to repeat their ratings shows that a numerical system of values given to the chosen features of the bone scans is reproducible. Furthermore, the good agreement between the rating of the three scan readers is reliable, since the ratings were done separately and independently by each evaluator and each looked at a different sequence of the scans. Two additional readers, one being a nuclear medicine technologist, also rated part or all of the scans and their results were virtually the same as those obtained by the three principal scan readers.

The diversity of rating values obtained for various patients with each of the agents used indicates that a valid comparison of several agents can be accomplished only if scans from a sufficient number of subjects are employed. Statistical analysis indicates that good results are possible if care is taken to standardize preparation techniques, randomize the sequence of usage, and maintain objectivity in the evaluations. Comparison of all of the agents in the same series of patients would be desirable, but some aspects of this type of study that may have a pronounced effect on the results have not been investigated. For example, a recent article (8) has indicated
that prior dosage with a polyphosphate, pyrophosphate, or diphosphonate bone-visualizing agent subsequently leads to unusual uptake of pertechnetate in several sites in the body. To date there is no evidence in the literature to show under what conditions initial dosing with one Tc-99m phosphate radiopharmaceutical would not affect the bone-uptake properties of another (or the same) technetium-labeled phosphate agent subsequently administered. Although one article (1) did report studies using alternate dosing of two agents in a series of patients, insufficient data were presented to determine whether any interaction between the first and second agents had occurred. Furthermore, if interaction could and did occur, factors such as the time lapse between the first and second doses, as well as the amount of chemical administered, may alter the degree of interaction. Until these factors are studied and any interaction between the agents or doses has been negated, the type of study we have presented is probably the most reliable.

In our work, no attempt was made to evaluate quantitatively any agent with respect to its degree of uptake in abnormal sites of bone metabolism, nor were the agents evaluated with respect to their efficiencies in diagnosing abnormalities. Nevertheless, inspection of the scans did seem to indicate that the sites of abnormal uptake were generally more readily perceived with those agents that received a high rating in this study.

ACKNOWLEDGMENTS

We thank Richard Larsen of the Mathematics Department of Vanderbilt University for his assistance in the statistical treatment of the results. We also appreciate the help of Robert Bowen and Paul A. Silk of the Nuclear Medicine Service of the Veterans Administration Hospital, Nashville, Tennessee for furnishing the scans and helping us to set up the evaluation criteria.

FOOTNOTES

* Procter and Gamble, Cincinnati, Ohio.
† Mallinckrodt Inc., St. Louis, Mo.
‡ Diagnostic Isotopes, Upper Saddle River, N.J.

REFERENCES

Clinical Evaluation of Some Phosphorus Bone-Imaging Agents: Concise Communication

Manno F. Nelson, L. Clifford McKee and John R. Van Wazer