Medical Observation, Compared with Parathyroidectomy, for Asymptomatic Primary Hyperparathyroidism: A Prospective, Randomized Trial

Jens Bollerslev, Svante Jansson, Charlotte L. Mollerup, Jörgen Nordenström, Eva Lundgren, Ove Tørring, Jan-Erik Varhaug, Marek Baranowski, Sylvi Aanderud, Celina Franco, Bo Freyschuss, Gunhild A. Isaksen, Thor Ueland, and Thord Rosen, on behalf of the SIPH Study Group*

Section of Endocrinology (J.B., G.A.I., T.U.), Rikshospitalet-Radiumhospital Medical Center, University of Oslo, N-0027 Oslo, Norway; Departments of Surgery (S.A.) and Endocrinology (C.F., T.R.), Sahlgrenska Hospital, S-416 85 Gothenburg, Sweden; Department of Endocrine Surgery (C.L.M.), Copenhagen University Hospital, DK-2100 Rigshospitalet, Denmark; Department of Surgery (J.N.), Karolinska University Hospital Solna, SE-171 76 Stockholm, Sweden; Department of Surgery (E.L.), Akademiska Hospital, SE-751 85 Uppsala, Sweden; Section of Endocrinology (O.T.), Department of Medicine, Södersjukhuset, SE-118 83 Stockholm, Sweden; Departments of Surgery (J.-E.V.) and Endocrinology (S.A.), Haukeland University Hospital and Institute of Medicine, University of Bergen, N-5021 Bergen, Norway; Department of Medicine (M.B.), St. Olav’s Hospital, University of Trondheim, 7006 Trondheim, Norway; and Department of Metabolism and Endocrinology (B.F.), Karolinska University Hospital, SE-141 86 Huddinge, Sweden

Objective: The present study is a randomized, controlled trial that investigates the effects of parathyroideectomy or medical observation in mild asymptomatic pHPT on morbidity and quality of life (QoL).

Design/Setting/Patients: A total of 191 patients (26 men) with asymptomatic pHPT (mean age 64.2 ± 7.4 (SD) yr) were recruited in the study and randomized to medical observation (serum calcium level 2.69 ± 0.08 mmol/liter) or surgery (2.70 ± 0.08 mmol/liter). We here report baseline and 1 (n = 119) and 2 yr data (n = 99) on those who had completed the follow-up visits by the end of the inclusion period.

Context: The clinical presentation of primary hyperparathyroidism (pHPT) has changed dramatically as a result of the increasing accessibility to biochemical analyses over the last decades. The introduction of biochemical autoanalyzers was followed by a rise and later a fall in the incidence of pHPT (1, 2). Today the diagnosis in the majority of cases is made accidentally in patients with mildly increased serum calcium and few or no specific symptoms. Whereas operative treatment in patients with higher calcium levels or organ-specific symptoms is well established, the value of surgical treatment of mild or borderline pHPT in asymptomatic patients has been a matter of discussion. So far, two consensus development conferences have been performed to provide guidelines and perspectives for treatment in relation to the change in the clinical presentation of the disease (3, 4). In the absence of randomized, controlled trials, these recommendations were based on clinical experience and data derived from epidemiological studies or nonrandomized, prospective cohort studies.

Results: At baseline, the patients had significantly lower QoL (SF-36) and more psychological symptoms, compared with age- and sex-matched healthy subjects. The two groups were similar at baseline, and no clinically significant changes in these parameters were seen during the observation time. Calcium and PTH normalized after surgery. The areal bone mineral density increased in the group randomized to operation, whereas the bone mineral density remained stable in the medical observation group. No change in kidney function (creatinine) or blood pressure was observed longitudinally or between the groups.

Conclusions: Asymptomatic patients with mild pHPT have decreased QoL and more psychological symptoms than normal controls. No benefit of operative treatment, compared with medical observation, was found on these measures so far. (J Clin Endocrinol Metab 92: 1687–1692, 2007)
and randomized by block randomization to medical observation or parathyroidectomy by an experienced parathyroid surgeon.

The diagnostic evaluation was based on local tradition of the single center and local laboratory. The diagnosis was based on elevated fasting serum calcium values on 3 occasional days corrected for variation in albumin levels. Moreover, at diagnosis serum level of intact PTH was measured at least twice, and both values should be above the mean of the reference interval at the local laboratory. Thiazide diuretics were stopped at least 4 wk before diagnostic evaluation. Although medication interfering with calcium metabolism was not accepted in the primary protocol, it was early decided to allow estrogens and bisphosphonates if the medication was not due to parathyroid bone disease. Thus, at study entry a total of 28 patients were on such medication. In the medical observation group, nine patients received estrogens and three bisphosphonates; in the operation group, 14 were on estrogens and two on bisphosphonates.

Patients stratified to medical observation were seen 3 months after randomization for safety reasons and then yearly. If conservatively followed patients developed symptoms or indications for surgery or demanded surgery, they were offered surgery. By the end of the inclusion period, a total of 10 patients randomized to medical observation have so far been surgically treated. Of these, seven had surgery due to medical reasons (four because of increasing calcium levels, one had diagnosed Paget’s disease, two because of poor general health), three at the request of the patients. However, in five the operation was performed after 2 yr of observation, and in only two was the operation performed during the 2 yr of observation. In the statistical analyses, they were regarded as medical observation patients (Intention to Treat). Patients randomized to operation were seen 3 months after surgery for safety reasons and then once yearly. Complications of surgery, e.g., hypocalcemia, were treated according to local traditions. In the case of unsuccessful primary operation, a secondary operation could be offered according to the protocol. However, no patients have been operated on more than once.

**Biochemistry and bone mineral density (BMD)**

Serum calcium, albumin, creatinine, and intact PTH (1-84) were measured by standard laboratory methods at the local laboratories. Over the years serum calcium was corrected for variation in normal reference by regression. Ionized calcium has been used at a single center (Copenhagen) and these values have been transformed to total values by regression. Areal BMD at the lumbar spine, femoral neck, and forearm was measured by dual-energy x-ray absorptiometry. Most centers used DPX-L software (version 1.31; Lunar Corp., Madison, WI). The center in Trondheim used Hologic QDR-4500 (Waltham, MA) in dual-beam mode. The center in Copenhagen used Norland XR 46 system 8091 (Fl. Atkinson, WI). BMD values in the lumbar spine and femoral neck were normalized as described (16).

**QoL assessments**

The primary psychosocial questionnaire used was the generic Short Form-36 general health survey (SF-36) (17), in which higher scores in-

---

**Subjects and Methods**

**Subjects**

From 1999 to the end of inclusion by the end of June 2005, a total of 191 patients (26 men) with asymptomatic pHPT were included in the study. Of the total population, 10 patients were included in Denmark, 126 in Sweden, and 55 in Norway. The mean age at inclusion for the study. Of the total population, 10 patients were included in Denmark, 126 in Sweden, and 55 in Norway. The mean age at inclusion for the end of the inclusion period, 24 patients had withdrawn (12 in the medical observation group); one was excluded due to parathyroid carcinoma. During the study nine patients decided to withdraw for other medical reasons (four because of increasing calcium levels, one had diagnosed Paget’s disease, two because of poor general health), three at the request of the patients. However, in five the operation was performed after 2 yr of observation, and in only two was the operation performed during the 2 yr of observation. In the statistical analyses, they were regarded as medical observation patients (Intention to Treat). Patients randomized to operation were seen 3 months after surgery for safety reasons and then once yearly. Complications of surgery, e.g., hypocalcemia, were treated according to local traditions. In the case of unsuccessful primary operation, a secondary operation could be offered according to the protocol. However, no patients have been operated on more than once.

**TABLE 1. Selection criteria for participation in the study**

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated, asymptomatic pHPT</td>
<td>Hyperparathyroid bone disease</td>
</tr>
<tr>
<td>2.60 ≤ S-Ca ≤ 2.85 mmol/liter</td>
<td>Previous operation in the neck</td>
</tr>
<tr>
<td>Age between 50 and 80 yr</td>
<td>Impaired kidney function (creatinine level &gt; 130 μmol/liter)</td>
</tr>
<tr>
<td>No medications interfering with Ca metabolism</td>
<td>Kidney stones</td>
</tr>
<tr>
<td>Informed consent</td>
<td>Complicating medical conditions</td>
</tr>
<tr>
<td></td>
<td>Psychiatric disorders</td>
</tr>
<tr>
<td></td>
<td>Multiple endocrine neoplasia, familial hypocalcemic hypercalcemia, familial HPT</td>
</tr>
</tbody>
</table>

S-Ca, Serum calcium.
dicate a better QoL. Because most of the patients in the study are from Sweden and there in general is only minor socioeconomic difference between the Scandinavian countries, the baseline data for the whole population was compared with normative Swedish data (18). An age- and sex-matched reference sample (n = 404) was randomly drawn from the Swedish SF-36 national normative database (n = 8950). A modification of the Comprehensive Psychopathological Rating Scale (CPRS) (15, 19) was used to give detailed information on mental symptoms. Data from this test were compared with normal Swedish control from a prospectively followed cohort (15).

Statistics

For analyzing treatment effects on paired data, univariate repeated-measures ANOVA were performed a priori with time as fixed factors and subjects as random. Data not normally distributed at baseline as evaluated by the Kolmogorov-Smirnov test were logistically transformed before inclusion in the general linear model (serum PTH and proximal radius BMD). Paired or unpaired t-tests for two samples were used when appropriate. Data from the SF-36, CPRS, and Z-scores for BMD were analyzed by nonparametric tests (Wilcoxon signed-rank test for paired data, Mann-Whitney U for unpaired data). Data are given as mean ± sd. P values are two sided and considered significant when <0.05.

Results

Baseline data

Baseline characteristics for the patient population are given in Table 2. A total of 95 patients were randomized to medical observation and 96 to surgery. As given in the table, no differences in biochemistry or bone mass could be demonstrated at the time of entry in study. It is also shown that the hypercalcemia was mild with a mean albumin corrected calcium levels of 2.69 and 2.70 mmol/liter in the two groups, respectively. Compared with normative data, BMD was not different from expected values when assessing Z-scores in the different compartments (data not shown).

Table 3 gives the baseline SF-36 score, compared with Swedish normative data. The patient population scored significantly lower in the psychological domains and in mental component summary score, compared with normal. Also in the comprehensive psychological rating scale, the patient population scored significantly worse, indicating more psychiatric symptoms than the matched control population.

Concerning mental health, no differences between the two groups were demonstrated at baseline (Figs. 1 and 2).

Analyzing baseline data for women alone or excluding patients on estrogens or bisphosphonates did not give different results concerning biochemistry, bone mass, or QoL.

Longitudinal data

The longitudinal observations on biochemistry and bone mass are given in Table 2. As seen, no significant changes in biochemistry were observed in the nonsurgery group after 2 yr of observation. Specifically, there were no changes in

### Table 3. Baseline QoL score (SF-36) and CPRS both in relation to normative Swedish data

<table>
<thead>
<tr>
<th>Description</th>
<th>Patients</th>
<th>Normal values</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF-36 standardized scores</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physiological functioning</td>
<td>74.9 ± 22.1</td>
<td>75.1 ± 26.1</td>
</tr>
<tr>
<td>Role physical</td>
<td>66.3 ± 40.5</td>
<td>72.5 ± 38.6</td>
</tr>
<tr>
<td>Bodily pain</td>
<td>66.8 ± 29.0</td>
<td>68.0 ± 28.7</td>
</tr>
<tr>
<td>General health</td>
<td>66.4 ± 22.2</td>
<td>68.9 ± 23.7</td>
</tr>
<tr>
<td>Vitality</td>
<td>57.5 ± 26.4</td>
<td>67.7 ± 25.4</td>
</tr>
<tr>
<td>Social functioning</td>
<td>83.2 ± 23.3</td>
<td>86.7 ± 22.0</td>
</tr>
<tr>
<td>Role emotional</td>
<td>67.8 ± 41.1</td>
<td>80.9 ± 33.6</td>
</tr>
<tr>
<td>Mental health</td>
<td>74.6 ± 21.4</td>
<td>80.9 ± 33.6</td>
</tr>
<tr>
<td>Physical component summary</td>
<td>44.6 ± 11.0</td>
<td>45.2 ± 11.6</td>
</tr>
<tr>
<td>Mental component summary</td>
<td>46.3 ± 12.4</td>
<td>51.0 ± 10.4</td>
</tr>
<tr>
<td>CPRS</td>
<td>9.63 ± 7.63</td>
<td>6.22 ± 4.10</td>
</tr>
</tbody>
</table>

Data are presented as mean and sd.

α P < 0.05.

β P < 0.01.

### Table 2. Biochemistry and BMD in 191 patients randomized to observation and operation at baseline and during 2 yr of incomplete follow-up

<table>
<thead>
<tr>
<th>Description</th>
<th>Baseline</th>
<th>1 yr</th>
<th>2 yr</th>
<th>P value for treatment × time effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observation (n = 95)</td>
<td>Surgery (n = 96)</td>
<td>Observation (n = 95)</td>
<td>Surgery (n = 96)</td>
</tr>
<tr>
<td>Biochemistry</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sPTH (pmol/liter)</td>
<td>10.19 ± 4.07</td>
<td>10.51 ± 4.07</td>
<td>11.40 ± 6.19</td>
<td>4.76 ± 2.17&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Albumin-corrected S-Ca (mmol/liter)</td>
<td>2.69 ± 0.08</td>
<td>2.70 ± 0.08</td>
<td>2.68 ± 0.13</td>
<td>2.40 ± 0.14&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Creatinine (μmol/liter)</td>
<td>80.8 ± 13.8</td>
<td>82.0 ± 13.7</td>
<td>81.09 ± 16.46</td>
<td>83.50 ± 19.13</td>
</tr>
<tr>
<td>Bone mass (BMD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LS (g/cm&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>1.06 ± 0.19</td>
<td>1.09 ± 0.19</td>
<td>1.04 ± 0.20</td>
<td>1.13 ± 0.19&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>FX (g/cm&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>0.83 ± 0.13</td>
<td>0.85 ± 0.14</td>
<td>0.81 ± 0.12</td>
<td>0.87 ± 0.13</td>
</tr>
<tr>
<td>Proximal radius (g/cm&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>0.58 ± 0.14</td>
<td>0.58 ± 0.15</td>
<td>0.56 ± 0.13</td>
<td>0.60 ± 0.16</td>
</tr>
<tr>
<td>Ultradistal radius (g/cm&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>0.34 ± 0.15</td>
<td>0.35 ± 0.15</td>
<td>0.31 ± 0.14</td>
<td>0.36 ± 0.17</td>
</tr>
<tr>
<td>MAP&lt;sup&gt;a&lt;/sup&gt; (mm Hg)</td>
<td>102.7 ± 12.2</td>
<td>103.0 ± 10.8</td>
<td>100.8 ± 11.5</td>
<td>102.9 ± 9.1</td>
</tr>
</tbody>
</table>

Data are presented as mean and sd. Only marginal differences were observed when patients not completing follow-up were excluded. The statistics are the same regardless, S-Ca, Serum calcium.

<sup>a</sup> P < 0.001 vs. observation group.

<sup>b</sup> P < 0.001 vs. baseline.

<sup>c</sup> P < 0.05 vs. observation group.

<sup>d</sup> P < 0.01 vs. baseline.

<sup>e</sup> P < 0.01 vs. observation group.

<sup>f</sup> Mean arterial pressure.
calcium or PTH levels. In the surgery group, parathyroidectomy was followed by normalization of calcium and PTH levels and remained stable from yr 1 to yr 2. Concerning serum creatinine levels, no changes were observed in any of the groups over time. Again, no different results were found when analyzing women alone or excluding patients on bone active medications at entry.

A significant increase in BMD was found for the lumbar spine (LS) 1 and 2 yr after surgery, compared with baseline as well compared with the medical observation group (Table 2). A similar trend was seen for the femoral neck (FN); however, the ANOVA analysis was of borderline significance only (P = 0.071). When analyzing women separately, this trend became significant (P = 0.048). Excluding patients on estrogens made the ANOVA analyses of borderline significance for both compartments (P = 0.080 for LS and P = 0.099 for FN). Excluding patients on bisphosphonates gave the same ANOVA results as for the whole group (P = 0.024 and P = 0.099 for LS and FN, respectively). No significant changes were observed for the forearm measurements in any of the groups or subgroups or in the nonsurgery group for any compartment.

The longitudinal QoL data are presented in Fig. 1. For SF-36 only minor and inconsistent findings were observed. Concerning the physical domains (Fig. 1, upper panel), a slightly, but significant, decrease was observed over the 2-yr period in the medical observation group (P < 0.01), whereas no change was seen in the operation group. The difference over time (delta values, not shown) was significantly different in favor of surgery (P < 0.01). As given in the lower panel of Fig. 1, the operation group scored slightly higher at year 1, compared with baseline in the mental health subdomain and mental component summary score (P < 0.05 for both), but not after 2 yr of observation. For the mental health subdomain, the observation group scored higher at 2 yr, compared with baseline (P < 0.05). Although no longitudinal differences were observed in any group in the other psy-
chological domains, the differences over time (delta values, not shown) for the domain role emotional were in favor of surgery for both 1 and 2 yr of observation.

The follow-up data for the CPRS questionnaire, compared with the controls, are given in Fig. 2. Throughout the observation period, the patient populations had significantly more symptoms than controls, also followed longitudinally. Only minor changes were observed within the groups and no significant differences in delta values (not shown) were observed, although a trend toward a better outcome in the operation group, compared with the observation group at yr 1 ($P = 0.079$). This is also seen in the longitudinal analysis within the operation group, which in the paired analysis scored significantly better after 1 yr, compared with baseline ($P < 0.05$), with the same trend after 2 yr of observation, but insignificant (Fig. 2). The longitudinal data on QoL and psychological symptoms were in principle not different when analyzing women alone or the subgroups without bone active medication at baseline.

**Discussion**

The present study shows that patients with marginally elevated calcium levels due to mild pHPT, who are quoted to be asymptomatic, do have decreased QoL and more psychological symptoms, than normal controls. For a total of 317 patient-years of follow-up, after randomization to either surgical treatment or medical observation, no systematic alterations were observed in either of treatment groups concerning SF-36 or the CPRS score. Neither could any differences in QoL and psychological symptoms of clinical importance be detected between the groups. The study confirms that medical observation is safe when patients are closely followed up (12) because the calcium-related biochemistry remained stable, as did creatinine and blood pressure. So far, only a few patients in the observation group needed to be operated on due to medical reasons. The study demonstrates that bone mass does not deteriorate during observation. Parathyroidectomy was followed by a decrease in PTH levels and thereby the expected increase in BMD of the lumbar spine (and femoral neck) due to closure of the remodeling space (21).

Decreased QoL and psychological symptoms have been discussed as indication for parathyroidectomy in pHPT, and it has been questioned whether the current recommendations for surgery are too conservative (22, 23). To our knowledge, only one single trial has addressed longitudinal changes in QoL in a randomized study of patients with mild pHPT (24) by means of the general health-related QoL questionnaire SF-36, showing a modest measurable benefit of operation, compared with medical observation for the domain role emotional, as was also observed in the present study. These data are in alignment with nonrandomized studies of QoL in pHPT (14, 22). We also noted a significant difference in the physical component summary after 2 yr observation in favor of operation. However, these modest and minor changes are of questionable clinical significance (25, 26) and should be regarded with caution because the questionnaires used were not disease specific and as such to our knowledge has not been developed.

As in previous studies of mild pHPT, we found that serum calcium and PTH levels or mean arterial blood pressure do not deteriorate during medical observation (12, 24, 27). At the entry of study, the patient population in the present study had normal bone mass in the various compartments when compared with normative Z-scores. Cancellous bone is preserved in pHPT, whereas there seems to be an obligatory cortical bone loss, at least in the forearm (21, 28–30). The unaltered bone mass in the different compartments in the observation group is in accordance with previous studies on mild pHPT (12, 24). In this study, we could demonstrate a significant increase in BMD only after surgery in the LS, although the BMD at the FN tended to increase. The increase in bone mass is explained by a decrease in bone turnover after the fall in PTH and thereby a closure of the remodeling space (21, 29). As found by others (12, 24, 31) with an observation period up to 17 yr after parathyroidectomy, we could not demonstrate a positive effect on the critical bone mass in the forearm of patients with pHPT. Epidemiological case-control studies have demonstrated an increase in fracture rate, particularly in the forearm (20, 32), which seems to decline after surgery. However, it is still unclear whether surgery will reduce the risk of fracture, compared with medical observation, in a prospective, randomized setting of mild pHPT. Because a few patients were on antosteoporotic drugs in our study, the information on fracture risk may not be interpretable. However, from a point of generalizability, it is more appropriate to compare outcomes in a true clinical setting, allowing the subjects to be optimally treated in all aspects.

The strength of this study is clearly the prospective, randomized design; the close follow-up; and the adherence to the protocol. There are, however, also important limitations. First of all, neither the subjects nor the investigators were blinded to the treatment given. Therefore, it is not possible to rule out a placebo effect of surgery (22, 24), although the present study so far did not reveal significant positive effects of parathyroidectomy. It is a drawback that the study did not record potential patients not included in the study at the different centers. Thus, the strict criteria for participation in the study may have induced selection bias so that our population might not be representative. A few patients were on bone-active medications at entry. This might be of importance for all end points analyzed in this study. However, a sensitivity analysis excluding patients on estrogens or bisphosphonates did not result in a significant change of the results. Finally, several patients withdrew because they could not accept adherence to their randomization group. However, this was an initial problem only in the beginning of the study period and was solved by improved information.

In conclusion, asymptomatic patients with mild pHPT have decreased health-related QoL and more psychological symptoms than normal healthy controls. So far, however, we have not found any clinical significant benefit of operative treatment, compared with medical observation. Biochemistry, blood pressure, and bone mass seem to be stable during medical observation, whereas parathyroidectomy is followed by a normalization of calcium and PTH and thereby
an increase in BMD in the spine and probably in the femoral neck.

Acknowledgments

The SIPH Study Group was established in 1994. Besides the authors of this paper, the following have contributed significantly to the project: Göran Åkerström, M.D., Ph.D., Akademiska Sjukhuset, Uppsala, Sweden; Lars Ove Farnebo, M.D., Ph.D., Karolinska Institute, Stockholm, Sweden; Leif Mosekilde, M.D., Ph.D., Århus Amtssygehus, Århus Denmark; Jan Falch, M.D., Ph.D., Aker University Hospital, Norway; Per Pfeffer, M.D., Ph.D., National University Hospital, Oslo, Norway; Catharina Stjøberg, Karolinska University Hospital, Huddinge, Stockholm, is thanked for technical assistance.

Received August 22, 2006. Accepted January 26, 2007.

Address all correspondence and requests for reprints to: Jens Bollerslev, M.D., D.Med.Sc., Section of Endocrinology, Department of Medicine, National University Hospital, N-0027 Oslo, Norway. E-mail: jens.bollerslev@medisin.uio.no.

This work was supported by the Norwegian Research Council. Disclosure Statement: S.J., C.L.M., E.L., M.B., S.A., C.F., B.F., G.A.I., and Biovitrum. J.B. received lecture fees from Novartis, Novo Nordisk, and T.U. have nothing to declare. J.N. received lecture fees from Amgen and Biovitrum. E.L. received lecture fees from Nycomed and Amgen.

References


JCEM is published monthly by The Endocrine Society (http://www.endo-society.org), the foremost professional society serving the endocrine community.