Longitudinal study of depot medroxyprogesterone acetate (Depo-Provera®) effects on bone health in adolescents: study design, population characteristics and baseline bone mineral density

Christine C. Johnsona,⁎, Ronald T. Burkmanb, Melanie A. Goldc, Robert T. Brownd,1, Zeev Harelc, Ann Brunerf, Margaret Stagerg, Laura K. Bachrachh, S. Paige Hertweckil, Anita L. Nelsonj, Dorothy A. Nelsonk, Susan M. Coupeyl, Alison McLeodm, Henry G. Bone

aHenry Ford Hospital, Detroit, MI 48202-3450, USA
bBaystate Medical Center, Springfield, MA 01199, USA
cChildren's Hospital of Pittsburgh, Pittsburgh, PA 15213, USA
dThe Ohio State University, Children's Hospital, Columbus, OH 43205, USA
eHasbro Children's Hospital and Brown University, Providence, RI 02903, USA
fJohns Hopkins University School of Medicine, Baltimore, MD 21287, USA
gMetro-Health Medical Center, Cleveland, OH 44109-1998, USA
hStanford University School of Medicine, Stanford, CA 94305-5208, USA
iUniversity of Louisville, Louisville, KY 40202, USA
jDavid Geffen School of Medicine at UCLA, Harbor-UCLA Medical Center, Torrance, CA 90509-2910, USA
kWayne State University School of Medicine, Detroit, MI 48202, USA
lChildren's Hospital at Montefiore, New York, NY 10467, USA
mPfizer Ltd, Sandwich, Kent, CT13 9NJ, UK
nMichigan Bone and Mineral Clinic, Detroit, MI 48236, USA

Received 18 August 2007; revised 6 November 2007; accepted 6 November 2007

Abstract

Background: This analysis was conducted to assess the baseline data and design methodology within an observational longitudinal comparison of use vs. nonuse of the injectable (intramuscular) contraceptive depot medroxyprogesterone acetate (DMPA-IM) and its effect on bone mass in adolescent women.

Study Design: A prospective, observational, open-label, unmatched-cohort, safety study in females aged 11–18 years. Participants either self-selected DMPA-IM (Depo-Provera®) 150 mg to be administered every 12 weeks for up to 240 weeks with a 120-week post-treatment follow-up or were nonusers (users of nonhormonal contraception or sexually abstinent) who were to be followed up for up to 360 weeks. As each participant entered the study, bone mineral density (BMD) was measured by dual-energy X-ray absorptiometry at the lumbar spine, hip and femoral neck regions, along with total body bone mineral content; serum and urine specimens were obtained for assay of bone metabolism markers and participants’ histories of parity and tobacco and alcohol use were obtained.

Results: A total of 389 participants were enrolled: 169 elected to begin DMPA-IM; 26 chose nonhormonal methods and 194 were abstinent. The baseline characteristics indicated significant disparities between DMPA-IM users and nonusers: compared with the nonusers, DMPA-IM users had more advanced chronologic and gynecologic ages, were more likely to have smoked, been pregnant and included more blacks. These factors would likely influence bone accretion rates independent of DMPA-IM exposure. Comparison of participant BMDs with standard reference data revealed that the study cohorts did not match reference populations closely enough to make a direct between-cohort comparative analysis feasible.

Conclusions: The baseline differences in cohort characteristics preclude a meaningful comparison of mean BMD changes over time between DMPA-IM users and nonusers cohorts, and comparisons of changes in Z-scores between cohorts were also not appropriate. Therefore,
within-participant BMD decreases from baseline were established as safety thresholds, and the proportion of individuals crossing those thresholds on a persistent or progressive basis was identified as the revised primary end point.

© 2008 Elsevier Inc. All rights reserved.

Keywords: Adolescent; Bone density; Depot medroxyprogesterone acetate; Contraceptive agents

1. Introduction

Bone mass acquired during adolescence is a key determinant of adult skeletal health. Exposure to sex hormones during puberty plays a significant role in the accrual of bone mass, and sustained production of estrogens is required for the maintenance of bone mass in adolescent and adult women. Although the exact timing is a matter of some debate, it has been estimated that at least 90% of peak bone mass (PBM) is accrued by the age of 18 years, with approximately 40% being gained between the onset of sexual maturation (Tanner stage 2) and 18 years of age [1–3]. Decreased acquisition of bone mass during this time of accrual could affect the incidence of fragility fractures throughout life.

Historically, bone mineral density (BMD) has been used as an efficacy end point in many clinical trials of agents for the treatment or prevention of osteoporosis. BMD has also been used as a safety end point in clinical trials of agents in which an adverse effect on bone mass is anticipated. Clinical trials of both types have typically employed a randomized placebo- or active-control design in which a suitably comparable control group enables a quantitative determination of the effect of the test agent on bone mass.

The injectable (intramuscular) contraceptive depot medroxyprogesterone acetate (DMPA-IM; Depo-Provera®) has been used for many years as a highly effective long-term contraceptive. A number of longitudinal studies in adolescents have reported reduced BMD associated with use of DMPA-IM, as compared with nonuse, and compared with use of low-dose combination oral contraceptives [4–8]. A decrease in BMD would be expected, based on the hypogestrogenic effect of medroxyprogesterone acetate [9]. However, a cross-sectional evaluation of DMPA-IM use did not find a strong association between DMPA-IM use and BMD in women aged 14–18 years [10].

Available evidence suggests that loss of BMD, seen in association with DMPA-IM, is at least partly reversible as the serum estradiol level rises after discontinuation of DMPA-IM, with the extent of reversibility being dependent upon the length of the post-treatment observation period. Specifically, several studies have shown a substantial post-treatment recovery of BMD in adult women [11–15], and one study has shown that BMD recovery occurs in adolescents [16]. The effect of steroidal contraceptives (including DMPA) on fracture incidence, if any, cannot be determined based on the existing literature [17].

This prospective, long-term, observational clinical study was initiated to further investigate the effect of DMPA-IM on BMD in a sample of women aged 11–18 years. The study was conducted in 12 academic medical centers across the United States. Following the enrollment of all participants, a decision was made to convene a Data Safety Monitoring Board (DSMB, composed of R Burkman, C Johnson and H Bone) to review patient safety data more thoroughly on an individual basis, to monitor the safety of the entire participant population and to provide scientific oversight for a study solely focused on safety.

This report describes the demographic, behavioral, clinical and laboratory characteristics of the participants at baseline by their contraceptive choice and communicates the DSMB’s assessment of the impact of these characteristics on the original objectives of the study and the proposed data analysis plan.

The overall objective of this prospective, open-label, observational study was to evaluate the effect of DMPA-IM on bone mass in adolescent women. The original intention of the study was to compare the percent changes from baseline in BMD, measured by dual-energy X-ray absorptiometry (DXA), in a group of adolescent women who self-selected DMPA-IM 150 mg injections for contraception ("DMPA-IM users") with the corresponding changes seen in an unmatched group of adolescent women who either self-selected a nonhormonal form of contraception or were sexually abstinent ("nonusers" cohort). Lumbar spine (L1-L4) and total proximal femur (“total hip”) BMD were selected as the primary end points, while femoral neck BMD, whole body bone mineral content (BMC) and urine and serum markers of bone turnover [urinary N-telopeptide (u-NTx), bone-specific alkaline phosphatase (BAP) and osteocalcin (OC)] were included as secondary end points. Additional objectives of the study were to assess BMD changes that occurred over time after DMPA-IM use was discontinued and to compare them with the changes seen during parallel continued observation of the nonusers cohort.

The original objectives assumed (a) that the BMD changes seen in the nonusers cohort would reflect the “normal” or “expected” changes that occur during adolescence and (b) that the participants who elected DMPA-IM treatment would, in the absence of DMPA-IM, have followed this same “normal” pattern of bone accretion during adolescence, thereby allowing the treatment effect of DMPA-IM to be estimated by comparison. However, strategies that are commonly used to insure the direct comparability of two groups such as randomization of participants to treatment groups, or effective matching of participants and blinding to the study medications, were judged to be neither feasible nor appropriate in this clinical
setting (for an editorial discussion of these issues, see Gold and Bachrach [18]).

After reviewing the planned safety assessments and the baseline characteristics of the participants, the DSMB determined that the originally planned between-group comparisons could not be relied upon, for the reasons described below, and revision of the planned analytical approach was recommended.

2. Materials and methods

2.1. Study design

This was a nonrandomized, open-label, fixed-dose, prospective, unmatched observational cohort study in which female adolescents aged 12–18 years, who presented at the study clinics having self-selected their contraceptive method, were invited to participate. The institutional review boards of the participating medical institutions approved the protocol and all participants provided written informed consent as well as written parental consent if the participant was younger than 18 years of age. Key inclusion and exclusion criteria are summarized in Table 1. Per protocol, participants choosing DMPA-IM were to receive a 150-mg intramuscular injection every 12 weeks, for a treatment period of up to 240 weeks (20 injections), with an additional post-treatment follow-up period of 120 weeks. Participants who chose nonhormonal contraception or were sexually abstinent were to be followed up for up to 360 weeks.

2.2. Classification of participants

At baseline, eligible adolescent women requesting contraception at the study clinics and electing to initiate DMPA-IM for contraception were placed into the “DMPA-IM users” cohort. Those adolescents electing nonhormonal contraception were placed in the “nonusers” cohort. Women attending the same study clinics who were not sexually active were also recruited to the “nonusers” cohort. At any time during the study, participants were permitted to change their method of contraception, to begin using any Food and Drug Administration-approved contraceptive (including DMPA-IM), or to become sexually abstinent, with changes to be accounted for in the analyses.

2.3. Study evaluations

All study participants were scheduled to visit the clinic every 12 weeks (±7 days), with a maximum window of 4 weeks allowed around the 12-week visit interval. Baseline data collection included anthropometric measures, age at menarche, gynecologic age and sexual maturity rating. Details of family medical history (e.g., osteoporosis), previous pregnancies, smoking status, alcohol use and level of physical activity were obtained at screening by physician inquiry. Monitoring of serum and urine chemistry, BMD, whole body BMC and serum biochemical markers of bone remodeling was conducted at screening and at Weeks 24, 60, 84, 120, 144, 180, 204 and 240. Measurement of BMD and whole body BMC was also conducted at Week 300 and at Week 360 or the final visit.

BMD at the lumbar spine, total hip and femoral neck and whole body BMC were measured by DXA (Hologic, Waltham, MA, USA, or GE Lunar, Madison, WI, USA). During the study, Hologic spine and hip linearity phantoms were circulated among study sites, scanned without repositioning and the scans sent to the DXA quality assurance center (Synarc, Portland, OR, USA) to provide comparison data across the study equipment. To account for the different DXA equipment used in the study, BMD values obtained on GE Lunar instruments were converted to Hologic values using published standardization equations [19,20]. While recognizing that the conversion equations had not been validated for adolescents, this appeared to be the most practical means of making such a comparison. BMD Z-scores, defined as the number of standard deviations from the mean BMD for healthy age-matched female adolescents, were calculated using Hologic (Caucasian) reference data after calculation of individual BMD results [21]. BMD T-scores (race adjusted), defined as the number of standard deviations from mean BMD for healthy young adult women, were also calculated (Hologic provided the reference database used for calculation of lumbar spine T-scores; total hip and femoral neck reference data were from published sources [22,23]). Whole body BMC values were not converted; corresponding reference data for each manufacturer were used.

Serum and urine specimens used for estradiol and bone marker measurements were collected but not standardized for time of day or stage of the menstrual cycle. Serum estradiol levels were measured using a chemiluminescence microparticle immunoassay (Abbott ImX, Abbott Laboratories, Abbott Park, IL, USA). Markers of bone formation

| Table 1 |
| Key inclusion and exclusion criteria |
| Inclusion criteria* |
| • Aged 12–18 years |
| • Any menses in the 6 months before enrollment |
| • Negative pregnancy test |
| • Not presently lactating |

Exclusion criteria |
| • Spinal BMD Z-score at screening ≤−2 of matched young normal participants |
| • Previous DMPA-IM use in prior 6 months |
| • Significant concomitant medication (bone-modifying agents, glucocorticoids, heparin or anticonvulsants) |
| • Thromboembolic disease |
| • Pregnancy (known or suspected) |
| • Body weight ≥250 lbs (113.4 kg) |
| • Significant alcohol or substance abuse |

BMD, bone mineral density; DMPA-IM, intramuscular injection of depot medroxyprogesterone acetate.

* Specific criteria applicable in the case of postpartum women or those who had recently had an abortion.
(serum OC and BAP) and bone resorption (u-NTx) were centrally assayed (CRL-Medinet, Lenexa, KS, USA) by enzyme-linked immunoassay (OC and u-NTx) or immunoradiometry (BAP).

Dietary calcium intake was not obtained at baseline but was estimated from a questionnaire administered later in the study follow-up, after a protocol amendment. Income and education were assigned by zip code of residence, based upon median data obtained from the US population census conducted in 2000.

2.4. Safety monitoring

The primary role of the DSMB was to monitor individual and group changes in BMD from a safety perspective; however, because the study was exclusively focused on safety end points, the DSMB also provided scientific oversight of the study. General safety was monitored through adverse-event recording and by laboratory assays. The study investigators were blinded to results of participants’ BMD and bone-marker measurements unless a participant had BMD loss from baseline of ≥5% or Z-score less than −2.5. A BMD deficit of ≥5% vs. baseline on two consecutive follow-up visits would be classified as persistent loss, while a loss of ≥5% vs. baseline followed by a deficit of ≥8% would be classified as progressive loss. To screen for causes of secondary bone loss, any participant with significant bone loss compared with her baseline level had a serum sample analyzed for intact parathyroid hormone, 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D at the next visit and at the following visit, approximately 6 months later. The investigators and the DSMB also recognized that nutritional differences might have contributed to differences in baseline BMD independent of other variables. For that reason, beginning in 2003, study participants were asked to report calcium and vitamin intake from their diet and supplements using structured interview questions.

2.5. Data analysis

The original data analysis plan for the study, reviewed by the DSMB upon its formation, included between-group comparisons of changes in BMD and other study end points. In order to investigate the comparability of the two cohorts by DMPA-IM use vs. nonuse, the DSMB requested a number of additional analyses of the baseline data, including comparison of demographics and personal characteristics; BMD Z-score (BMD referenced to gender- and age-specific bone mass) and BMD T-score (BMD referenced to race- and gender-specific PBM) according to age, cohort, race and DXA manufacturer. It was fully recognized that T-scores are not appropriate for the classification of adolescent patients in clinical practice, but they were considered as part of the overall characterization of the study groups.

Baseline data were summarized by DMPA-IM users vs. nonusers, including stratified analyses of the nonusers cohort (nonhormonal methods vs. abstinent). Continuous variables were analyzed by one-way analysis of variance (ANOVA), and categorical variables were compared using chi-squared tests. To account for skewed data, alkaline phosphatase and estradiol values were log-transformed before being modeled using a one-way ANOVA. Statistical tests were two-sided and a statistical test with a p value of <.05 was considered statistically significant. Associations between personal characteristics at baseline and baseline BMD were assessed using stepwise regression. All statistical analyses were conducted using SAS® release 8.2 (SAS Institute, Cary, NC, USA).

3. Results

The analysis of the baseline characteristics of the study population revealed notable differences between the cohorts of adolescent women who had made different contraceptive choices. A total of 389 participants were enrolled at 12 study centers. An additional 26 participants were enrolled at a 13th center, but that center was closed shortly after the participants enrolled and no follow-up data were obtained for these participants. Assessment of baseline data including and excluding data from the closed site showed no statistically significant differences and no baseline data for participants from the closed center are reported here. Of the 389 adolescents who participated in the study, 169 chose DMPA-IM contraception (“DMPA-IM users”), and 220 did not choose DMPA-IM (“nonusers”; 26 elected nonhormonal forms of contraception, almost exclusively condoms; 194 were sexually abstinent). Baseline characteristics are shown in Table 2A and B. In light of the heterogeneity of the nonusers cohort, baseline characteristics were also analyzed separately for participants using nonhormonal methods (“nonhormonal” subgroup) and those who were sexually abstinent (“abstinent” subgroup).

The mean chronologic age differed significantly between the DMPA-IM users and nonusers cohorts (p<.001), with an excess of younger participants entered into the abstinent cohort (Table 2C). Further analysis of the nonusers cohort showed that, on average, the participants who chose nonhormonal methods were older and the abstinent population was younger than those who chose DMPA-IM. While age at menarche was not different among the three cohorts, gynecologic age at the time of study entry mirrored chronologic age and also differed significantly between the cohorts (mean 44.4, 59.4 and 33.4 months post menarche in the DMPA-IM users, nonhormonal and abstinent cohorts, respectively; p<.001). The racial composition of the three cohorts was also different (p<.001). The DMPA-IM users cohort was 60% black (101/169), while only 33% (64/194) of abstinent participants were black; overall, 36% (79/220) of the nonusers were black (Table 2A).

The majority of participants in the DMPA-IM users and nonhormonal cohorts — but only a small proportion of the abstinent cohort — had previously used contraception. The
percentages of participants who smoked cigarettes, consumed alcohol or had a family history of breast or urogenital cancer or cardiovascular disease were lower in the abstinent cohort (Table 2B). None of 194 participants in the abstinent cohort had ever been pregnant; in contrast, 27% (7/26) of the nonhormonal cohort and 17% (29/169) of the DMPA-IM users reported a previous pregnancy (p<.001; Table 2B). With respect to race, prior use of contraception and tobacco or alcohol use, the profile of the small nonhormonal cohort was more similar to that of the DMPA-IM users.

Socioeconomic analysis, which was based solely on zip code of residence and used median data from the 2000 US

Table 2A  
Population demographics at baseline

<table>
<thead>
<tr>
<th>Baseline characteristics</th>
<th>DMPA-IM users (n=169)</th>
<th>Nonusers (n=220)</th>
<th>p value&lt;sup&gt;bc&lt;/sup&gt;</th>
<th>Nonusers Nonhormonal (n=26)</th>
<th>Abstinent (n=194)</th>
<th>p value&lt;sup&gt;bc&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronologic age (years), mean (S.D.)</td>
<td>15.4 (1.6)</td>
<td>14.6 (1.7)</td>
<td>&lt;.001</td>
<td>16.3 (1.5)</td>
<td>14.3 (1.5)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age at menarche (years), mean (S.D.)</td>
<td>11.7 (1.6)</td>
<td>11.5 (1.4)</td>
<td>.210</td>
<td>11.5 (1.4)</td>
<td>11.5 (1.4)</td>
<td>.450</td>
</tr>
<tr>
<td>Gynecologic age (months), mean (S.D.)</td>
<td>44.4 (22.0)</td>
<td>36.5 (23.3)</td>
<td>&lt;.001</td>
<td>59.4 (24.0)</td>
<td>33.4 (21.5)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>101 (59.8)</td>
<td>79 (35.9)</td>
<td>15 (57.7)</td>
<td>64 (33.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>37 (21.9)</td>
<td>100 (45.5)</td>
<td>5 (19.2)</td>
<td>95 (49.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>31 (18.3)</td>
<td>41 (18.6)</td>
<td>6 (23.1)</td>
<td>35 (18.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian/Pacific islander</td>
<td>0</td>
<td>5 (2.3)</td>
<td>0</td>
<td>5 (2.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>21 (12.4)</td>
<td>27 (12.3)</td>
<td>3 (11.5)</td>
<td>24 (12.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>10 (5.9)</td>
<td>9 (4.1)</td>
<td>3 (11.5)</td>
<td>6 (3.1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reported previous use of contraceptive methods, n (%)  
Yes (n=176) 142 (84.0) 34 (15.5) 25 (96.2) 9 (4.6)  
No (n=212) 27 (16.0) 185 (84.1) 1 (3.8) 184 (94.8)  
Mean (S.D.) weight, kg  
61.66 (13.58) 61.24 (12.69) .751 65.64 (14.23) 60.65 (12.39) .179  
Mean (S.D.) BMI, kg/m²  
23.66 (4.82) 23.38 (4.39) .544 25.05 (5.12) 23.15 (4.25) .116

BMI, body mass index.  
a n=168 for weight and BMI.  
b Chi-squared test for race and previous use of contraception; other variables: one-way ANOVA.  
c For DMPA-IM users vs. nonusers.  
d n=192 for previous contraceptive use.  
e For between-cohort comparison: DMPA-IM users vs. nonhormonal vs. abstinent.  
f Although inclusion criteria specified the age range to be 12–18 years, one 11-year-old participant was enrolled.  
g For comparison of racial composition.

Table 2B  
Baseline characteristics (risk factors)

<table>
<thead>
<tr>
<th>Baseline characteristics</th>
<th>DMPA-IM users (n=169)</th>
<th>Nonusers (n=220)</th>
<th>p value&lt;sup&gt;bc&lt;/sup&gt;</th>
<th>Nonusers Nonhormonal (n=26)</th>
<th>Abstinent (n=194)</th>
<th>p value&lt;sup&gt;bc&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cigarette smoking and alcohol use, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>42 (24.9)</td>
<td>15 (6.8)</td>
<td>&lt;.001</td>
<td>10 (38.5)</td>
<td>5 (2.6)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Alcohol user</td>
<td>29 (17.2)</td>
<td>18 (8.2)</td>
<td>.007</td>
<td>9 (34.6)</td>
<td>9 (4.6)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Smoker and alcohol user</td>
<td>17 (10.1)</td>
<td>5 (2.3)</td>
<td>.001</td>
<td>3 (11.5)</td>
<td>2 (1.0)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Family medical history, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast or urogenital cancer</td>
<td>11 (6.5)</td>
<td>3 (1.4)</td>
<td>.007</td>
<td>0</td>
<td>3 (1.6)</td>
<td>.025</td>
</tr>
<tr>
<td>Osteoporosis or bone-related disorders</td>
<td>5 (3.0)</td>
<td>7 (3.2)</td>
<td>.895</td>
<td>2 (7.7)</td>
<td>5 (2.6)</td>
<td>.371</td>
</tr>
<tr>
<td>Diabetes</td>
<td>23 (13.7)</td>
<td>24 (11.0)</td>
<td>.425</td>
<td>4 (15.4)</td>
<td>20 (10.4)</td>
<td>.558</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>22 (13.1)</td>
<td>16 (7.3)</td>
<td>.060</td>
<td>4 (15.4)</td>
<td>12 (6.3)</td>
<td>.058</td>
</tr>
<tr>
<td>Previous pregnancies, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gravida ≥1</td>
<td>29 (17.2)</td>
<td>7 (3.2)</td>
<td>&lt;.001</td>
<td>7 (26.9)</td>
<td>0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Para ≥1</td>
<td>16 (9.5)</td>
<td>3 (1.4)</td>
<td>&lt;.001</td>
<td>3 (11.5)</td>
<td>0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Abortus ≥1</td>
<td>19 (11.2)</td>
<td>3 (1.4)</td>
<td>&lt;.001</td>
<td>3 (11.5)</td>
<td>0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Ectopic ≥1</td>
<td>0</td>
<td>1 (0.5)</td>
<td>NC</td>
<td>3 (1.8)</td>
<td>0</td>
<td>NC</td>
</tr>
</tbody>
</table>

NC, noncomputable (sparse data).  
a n=168 for family medical history.  
b n=218 for family medical history.  
c Chi-squared test.  
d For DMPA-IM users vs. nonusers.  
e n=192 for family medical history.  
f For between-cohort comparison: DMPA-IM users vs. nonhormonal vs. abstinent.
census, showed that participants in the DMPA-IM users cohort typically lived in locales with lower income and educational levels than participants in the nonusers cohort. The median household income for the areas inhabited by the nonusers cohort was 27% higher than for the DMPA-IM users cohort. Only 3% of DMPA-IM users lived in areas where the median household income was greater than $65,000, similar to the proportion of the nonhormonal cohort (4%), while 16% of abstinent participants lived in communities where the median income was above this level. Likewise, geographically assessed educational levels were lowest among DMPA-IM users.

The DSMB also addressed the question of comparability of the groups with respect to their bone mass and metabolism. There was a nonsignificant trend toward higher BMD at the lumbar spine and the femoral neck among DMPA-IM users, compared with all nonusers (Table 3A); the consistently lower baseline BMD values in the large abstinent cohort (n=194) appeared to be driving this trend. The significantly higher baseline BMD among black participants (Table 3B), along with the disproportionate enrollment of black participants in the DMPA-IM users cohort, appeared to be contributory factors. The relationship between baseline BMD and gynecologic age for DMPA-IM users and nonusers is shown for lumbar spine and for total hip in Fig. 1A and B, respectively. Comparison of the slopes of the regression lines for DMPA-IM users vs. nonusers suggests that the underlying relationship between baseline BMD and gynecologic age differs between the two cohorts. In addition, there is a significant disparity between the two groups in the proportion of participants less than 20 months post menarche [8% (14/169) for DMPA-IM users vs. 28% (62/219) for nonusers; p<.001, chi-squared test]. These younger participants would be expected to show the greatest increases in BMD during the study period, and hence, the proportion of younger participants would influence calculated group mean changes in BMD.

These findings were further explored by comparing participants’ baseline bone density data against available reference populations for Z-score and T-score (as detailed in Methods). As shown in Table 4, all cohorts at all ages had substantially higher baseline lumbar spine BMD Z-scores than the age-specific reference population. This pattern suggested that the study population at baseline was skeletally more mature than the reference population. Hence, it raised the possibility that subsequent rates of bone accretion during the study would be lower for study participants than for the

### Table 3A
Bone mineral density (standardized) and whole body bone mineral content at baseline

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DMPA-IM users (n=169)</th>
<th>Nonusers (n=219)</th>
<th>p value</th>
<th>Nonusers (n=219)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMD (g/cm²), mean (S.D.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lumbar spine</td>
<td>1.092 (0.113)</td>
<td>1.068 (0.137)</td>
<td>.071</td>
<td>1.106 (0.113)</td>
<td>.057</td>
</tr>
<tr>
<td>Total hip</td>
<td>1.028 (0.118)</td>
<td>1.019 (0.129)</td>
<td>.521</td>
<td>1.038 (0.140)</td>
<td>.591</td>
</tr>
<tr>
<td>Femoral neck</td>
<td>1.018 (0.125)</td>
<td>0.994 (0.125)</td>
<td>.059</td>
<td>1.009 (0.125)</td>
<td>.137</td>
</tr>
<tr>
<td>Whole body BMC (g), mean (S.D.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hologic</td>
<td>2095 (275.2)</td>
<td>2020 (367.2)</td>
<td>.808</td>
<td>2174 (219.4)</td>
<td>.103</td>
</tr>
<tr>
<td>Lumbar</td>
<td>2382 (332.1)</td>
<td>2392 (374.8)</td>
<td>.860</td>
<td>2505 (385.9)</td>
<td>.318</td>
</tr>
</tbody>
</table>

* Unless otherwise stated.
* One participant did not provide analyzable BMD data at baseline.
* One-way ANOVA.
* For DMPA-IM users vs. nonusers.
* For between-cohort comparison: DMPA-IM users vs. nonhormonal vs. abstinent.
reference population. Thus, the use of change in Z-scores was not appropriate to compare cohorts in this study. There were also observed differences in mean Z-scores between participants scanned with Hologic equipment, compared with those scanned using Lunar equipment (data not shown). This could have been a center effect or an indication that the published conversion equations introduced a systematic error in these young participants.

Assessment of T-scores is not recommended in clinical practice because adolescents are skeletally immature. Unlike postmenopausal women, in adolescents, there is no relationship between fracture risk and a BMD T-score less than −1.0 or −2.5, the traditional values used to define “osteopenia” or “osteoporosis,” respectively, in postmenopausal women. However, BMD T-scores were analyzed in this study as a proxy for absolute BMD, not only because T-scores provided

Fig. 1. Standardized bone mineral density at (A) lumbar spine and (B) total hip vs. gynecologic age for the study population at baseline.
an indicator of progress toward peak bone density (T-score=0) but also because T-scores allowed adjustment of the BMD data for race, an important factor in this study. When the race-adjusted baseline lumbar spine T-scores were viewed cross-sectionally according to chronologic age (Table 4), it became apparent that the underlying relationship between baseline T-score and age was different between the nonusers cohort and the DMPA-IM users cohort, again suggesting that the two groups were not comparable. Because the mean lumbar spine T-scores were, with few exceptions, greater than −0.5 (Table 4), these data also suggest that the adolescents in this study were already quite close to achieving PBM, with a substantial proportion of each group having T-scores above 0. Baseline T-scores for total hip and femoral neck showed less variation with age than was seen for lumbar spine (data not shown); across all participants, the mean baseline T-score was +0.31 for total hip and +0.91 for femoral neck. A total of nine participants had baseline lumbar spine T-scores less than −2.5 [1/169 (0.6%) in the DMPA-IM users cohort, 8/194 (4.1%) in the abstinent cohort]; none of the participants had a baseline total hip or femoral neck BMD T-score below −3.2.

Assay results for biochemical markers of bone remodeling (u-NTx, OC and BAP), as well as for estradiol, are shown in Table 5. Concentrations of all bone markers differed significantly across the three cohorts (p=0.031 to p<0.001 for DMPA-IM users vs. nonusers), with the highest median values seen in the abstinent cohort and the lowest values in the nonhormonal cohort, paralleling the between-group differences in mean gynecologic age. Fig. 2 shows baseline BMD vs. baseline urine N-telopeptide/creatinine ratio (u-NTx) for DMPA-IM users vs. nonusers. Values of u-NTx >200 nM bone collagen equivalents (BCE)/mM creatinine were seen in 6.2% of participants who chose DMPA-IM while 25.4% of nonusers had baseline u-NTx values in that range (p<0.002, chi-squared test), reflecting a higher level of bone turnover and indicative of relative skeletal immaturity. There were no statistically significant differences in estradiol concentrations between the cohorts (Table 5).

### 4. Discussion

The differences in baseline characteristics were reviewed by the DSMB during a series of meetings that resulted in a recommendation for changes in the planned analyses. It had originally been anticipated that study recruitment would provide relatively comparable cohorts in which the degree of

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DMPA-IM users</th>
<th>Nonusers</th>
<th>p value&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Nonusers</th>
<th>Abstinent</th>
<th>p value&lt;sup&gt;a, b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osteocalcin (ng/mL), median (range)</td>
<td>5.10 (0.3–29.5)</td>
<td>6.55 (0.7–36.6)</td>
<td>.031</td>
<td>4.75 (1.1–11.3)</td>
<td>6.65 (0.7–36.6)</td>
<td>.052</td>
</tr>
<tr>
<td>Urinary N-telopeptide</td>
<td>76.4 (14.5–620.0)</td>
<td>116.3 (10.8–892.0)</td>
<td>&lt;.001</td>
<td>66.7 (26.2–217.0)</td>
<td>120.0 (10.8–892.0)</td>
<td>.042</td>
</tr>
<tr>
<td>(nM BCE/mM creatinine, median (range))</td>
<td>(n=128)</td>
<td>(n=173)</td>
<td></td>
<td>(n=14)</td>
<td>(n=159)</td>
<td></td>
</tr>
<tr>
<td>Alkaline phosphatase (U/L), median (range)</td>
<td>31.15 (12.3–138.8)</td>
<td>38.30 (13.8–267.2)</td>
<td>&lt;.001</td>
<td>30.90 (18.5–46.2)</td>
<td>40.00 (13.8–267.2)</td>
<td>.007</td>
</tr>
<tr>
<td>(n=138)</td>
<td>(n=181)</td>
<td></td>
<td>(n=14)</td>
<td>(n=167)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estradiol (pg/mL), median (range)</td>
<td>74.0 (17–646)</td>
<td>80.0 (17–554)</td>
<td>.517</td>
<td>93.0 (35–289)</td>
<td>74.5 (17–554)</td>
<td>.714</td>
</tr>
<tr>
<td>(n=128)</td>
<td>(n=175)</td>
<td></td>
<td>(n=13)</td>
<td>(n=162)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Wilcoxon rank sum test (osteocalcin, N-telopeptide); one-way ANOVA after log transformation (alkaline phosphatase, estradiol).

<sup>b</sup> For between-cohort comparison: DMPA-IM users vs. nonhormonal users vs. abstinent.
confounding could be handled analytically and that the safety of DMPA-IM use in the study population could be assessed by comparison of BMD values with those of the nonusers. However, between-cohort differences in baseline data were sufficiently large to make comparisons between DMPA-IM users and nonusers inappropriate. In particular, the fundamental assumption of the study analysis plan — that the two groups would have shown similar BMD changes over time in the absence of any DMPA-IM being administered — was unlikely to be correct.

Serum samples for estradiol measurements at baseline were collected without regard to timing with respect to the participants’ menstrual cycles. The ranges were wide, as would be expected, and neither the ranges nor the medians differed materially between the cohorts. Thus, within the limitations of the sampling procedure, there was no indication that the groups differed with respect to estradiol levels prior to initiation of treatment. Therefore, it is unlikely that the between-group differences were attributable to the participants’ estrogen levels. This baseline similarity will be of importance in the interpretation of the response to hormone deprivation by treatment with DMPA-IM.

While the racial imbalance between the cohorts may be reflective of the contraceptive choices that adolescents make in the community [24], black adolescents tend to gain bone earlier during adolescence [25,26], and their rate of BMD accrual may plateau earlier. Each of these factors would contribute to a greater skeletal maturity at the onset of the study and would make it more likely that the DMPA-IM users would demonstrate smaller increases in BMD during follow-up, apart from any effect of DMPA-IM administration.

In addition, cross-sectional baseline BMD measurements for all cohorts were inconsistent with the age-specific (Z-score) reference values. For both cohorts, mean BMD at each age (in years) was higher than predicted by the Z-score reference population ($Z=0.0$ at each age). This was disproportionately evident at the lowest ages, and in the DMPA users cohort in particular, with the result that the increase in BMD across the age spectrum was much less pronounced in either study cohort compared with the Z-score reference population. The differences between the cohorts and between either cohort and the reference population suggested that (a) the study cohorts would not, even in the absence of DMPA-IM administration, accrue bone density at similar rates during the course of the study and (b) the originally planned comparisons based on changes in Z-scores, intended to demonstrate the specific effect of DMPA-IM on BMD, would be meaningless or misleading.

Based on its assessment of the baseline data, the DSMB recommended analyzing the study cohorts independently, without direct comparison. In order to address the concerns for ongoing safety assessment of bone health, the DSMB recommended a monitoring program to identify participants at particular risk for significant BMD decline and evaluate secondary causes of bone loss in those participants, as described above in Section 2.4 “Safety Monitoring”.

In summary, comparisons of baseline demographics, clinical characteristics and laboratory data from this study of DMPA-IM users and nonusers found significant differences between these self-selected cohorts at baseline. Chronologic and gynecologic ages were more advanced, and the proportion of black participants was greater, in the DMPA-IM users cohort. These factors would contribute to greater skeletal maturity in the DMPA-IM users cohort and would predict more modest subsequent gains in BMD in that group during the period of observation. That is, only part of

![Fig. 2. Standardized lumbar spine bone mineral density vs. urinary N-telopeptide level for the study population at baseline.](image-url)
any deficiency in BMD accretion seen in the DMPA-IM users cohort compared with the nonusers would be the result of DMPA-IM administration, with no means of quantifying that proportion. While the original objective of the study was to estimate the rate of change of BMD in the DMPA-IM users cohort and compare it with the rate of change for nonusers, the DSMB recognized that these differences could introduce uncontrollable confounding if the initially planned analyses were deployed. The DSMB concluded that valid between-cohort BMD comparisons could not be carried out because of these differences and that adequate adjustment procedures could not be devised that would allow such comparisons. Consequently, the DSMB decided that safety analyses for individual participants and cohort groups would be based on within-participant changes from baseline and that the cohorts would be assessed and reported separately. In accordance with the DSMB decision, these descriptive observational analyses are ongoing. These methodological considerations warrant emphasis in studies of BMD where nonrandomized observations are involved.

Acknowledgments

The authors would like to thank Pina D’Angelo and Carol Reid for their assistance with the statistical analyses, and Kevin Wolter (Pfizer) and Julie Ponting (Anthemis Consulting) for their assistance in the assembly and verification of data and preparation of the manuscript.

References


