DEPARTMENT OF HEALTH & HUMAN SERVICES

Food and Drug Administration
Rockville MD 20857

MAY 9 2006

Jeffrey G. Roberts
Sr. Vice President and Chief Technical Officer
Wright Medical Technology, Inc.
5677 Airline Road
Arlington, TN 38002

Re: Docket No. 2005P-0440

Dear Mr. Roberts:

This letter responds to your citizen petition (Petition) dated October 29, 2005, submitted on behalf of Wright Medical Technology, Inc. (Wright).[1] You ask (Petition at 1) that the Food and Drug Administration (FDA), pursuant to 21 U.S.C. 360e(d), deny approval of Smith & Nephew's (SN) premarket approval application (PMA) for the Birmingham Hip Resurfacing (BHR) system (P040033). Specifically, you ask that FDA determine that SN has not met its statutory burden of providing a reasonable assurance of safety and effectiveness. Id.

DECISION SUMMARY

Your request that we deny approval of SN's PMA is denied. Under the Federal Food, Drug, and Cosmetic Act (Act) and FDA regulations, the agency may approve PMAs where the data and information in support of the approval, when taken as a whole, provides reasonable assurance that the device is safe and effective for its conditions of use. We find your argument that the use of a "retrospective, uncontrolled case series at a single center by a single physician without any protocol and with incomplete follow-up" as not scientifically valid (March 10, 2006, letter at 16) to represent a mischaracterization of the sufficiency of the scientific evidence, unpersuasive, and contrary to the law.

OVERVIEW

In this petition response, we first explain that the agency, in accordance with the Act and FDA regulations, may approve a PMA based on various types of data and information, including data and information from a single investigation by a single investigator. Then we explain why your arguments are not persuasive. Section I sets forth the background, including relevant facts and statutory background. Section II sets forth the agency's analysis. Specifically, sections II.A-F discuss the data and information submitted in SN's PMA on which you expressed concerns. Section II.G explains that an approval based on the data and information submitted in the SN PMA is consistent with FDA's past practice. Section II.H discusses the agency's review process

[1] Prior to the submission of the Petition, you sent us a letter on October 11, 2005, which we responded to on December 2, 2005. In addition, comments to your Petition were submitted by Smith & Nephew on February 8, 2006, to which you responded on March 10, 2006. The agency's response to your petition is based on, among other things, a review of these submissions and the administrative record of the Smith & Nephew PMA.

for financial disclosure. Section III discusses the Advisory Panel's recommendations. Finally, section IV sets forth the agency's summary conclusion.

DISCUSSION

I. BACKGROUND

A. FACTUAL INFORMATION

SN submitted a PMA on July 19, 2004, under section 515(c) of the Act(21 U.S.C. 360e(c)) and 21 CFR 814.20, for the Birmingham Hip Resurfacing (BHR) system; a single use device intended for hybrid fixation, i.e., cemented femoral head component and cementless acetabular component. The BHR system is intended for use in patients requiring primary hip resurfacing arthroplasty due to: 1) non-inflammatory arthritis (degenerative joint disease) such as osteoarthritis, traumatic arthritis, avascular necrosis, or dysplasia/developmental dislocation of the hip (DDH); or 2) inflammatory arthritis such as rheumatoid arthritis. The BHR system is intended for patients who, due to their relatively younger age or increased activity level, may not be suitable for traditional total hip arthroplasty due to an increased possibility of requiring future ipsilateral hip joint revision. It conserves the femoral head, by resurfacing rather than removing it and therefore, unlike existing alternatives, does not require removing the patient's femoral head and several inches of femur. No other total hip replacement implant that conserves the femoral stock and head is currently legally marketed.

FDA referred SN's PMA to the Orthopaedic and Rehabilitation Advisory Panel (the Panel), under section 515(c)(3) of the Act (21 U.S.C. 360e(c)(3)), for a report and recommendation respecting the approval of the application. On September 8, 2005, SN presented the clinical data from the single investigator, Derek J.W. McMinn, FRCS[2] (Mr. McMinn) to the Panel, in addition to other data and information. The Panel discussed the safety and effectiveness of the device, the data gathering methods, and the applicability of the data to the United States population.[3] The Panel recommended, by a vote of three to two, that the BHR system should be approved with conditions. Subsequent to this Panel meeting, you submitted a letter, dated October 11, 2005, expressing various concerns related to what you described as an apparent inconsistency in the regulatory requirements for the SN PMA compared to Wright's CONSERVE Plus PMA. The agency responded to you on December 2, 2005. Prior to receiving the agency's response, you submitted the Petition on October 29, 2005.

---

[2] FRCS stands for the Fellowship of the Royal College of Surgeons of England. The FRCS diploma was awarded following success in an examination in the generality of surgery taken towards the end of basic surgical training and before entry to higher surgical training in England. Mr. McMinn is a qualified surgeon in the U.K. Information from the web site: http://www.rcseng.ac.uk/career/opportunities/internationaldoctors/glossary.html#f

[3] Information concerning this Panel meeting, including a transcript, can be found at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfAdvisory/details.cfm?mtg=600

B. RELEVANT STATUTORY AND REGULATORY BACKGROUND

1. Content of PMA

Under section 515(c) of the Act (21 U.S.C. 360e(c)), an applicant seeking to market a Class III device must first obtain FDA approval by filing a PMA for the device. PMAs must contain, among other things, scientific data demonstrating the safety and effectiveness of the device. Id. When making a decision to deny or approve a PMA, FDA is to rely on the conditions of use included in the proposed labeling as the basis for determining whether or not there is a reasonable assurance of safety and effectiveness (section 515(d)(1)(A) of the Act (21 U.S.C. 360e(d)(1)(A)).

The PMA, by regulation, must include a discussion demonstrating that the data and information in the application constitute valid scientific evidence, within the meaning of 21 CFR 860.7, and provide reasonable assurance that the device is safe and effective for its intended use (21 CFR 814.20(b)(3)(vi)). A PMA can be supported solely by data from one investigation, provided the application contains a justification showing that data and other information from a single investigator are sufficient to demonstrate the safety and effectiveness of the device and to ensure reproducibility of test results (21 CFR 814.20(b)(7)).

2. Valid Scientific Evidence

The agency relies upon only valid scientific evidence to determine whether there is reasonable assurance of safety and effectiveness and reviews the available evidence taken as a whole (21 CFR 860.7(c)(1)). Valid scientific evidence may include evidence from well-controlled investigations, partially controlled studies, studies and objective trials without matched controls, well-documented case histories conducted by qualified experts, and reports of significant human experience with a marketed device from which it can fairly and responsibly be concluded by qualified experts that there is reasonable assurance of safety and effectiveness of a device under its conditions of use (21 CFR 860.7(c)(2)). There is reasonable assurance that a device is safe where FDA can determine, based upon valid scientific evidence, that the probable benefits to health from use of the device for its intended uses and conditions of use, when accompanied by adequate directions and warnings against unsafe use, outweigh any probable risks (21 CFR 860.7(d)(1)). There is reasonable assurance that a device is effective where FDA can determine, based upon valid scientific evidence, that in a significant portion of the target population, the use of the device for its intended uses and conditions of use, when accompanied by adequate directions for use and warnings against unsafe use, will provide clinically significant results (21 CFR 860.7(e)(1)).

Congress provided the agency with broad discretion for how to determine whether there is reasonable assurance of safety and effectiveness for devices (see Ethicon v. FDA, 762 F. Supp. 382, 386 (D.D.C. 1991)). Under section 513(a)(3) of the Act (21 U.S.C. 360c(a)(3)), FDA may base a determination of effectiveness on valid scientific evidence other than well-controlled investigations when FDA determines such evidence is sufficient to determine the effectiveness of the device, and from which it can be fairly and responsibly concluded by qualified experts that the device will have the effect it purports or is represented to have under the conditions of use

prescribed, recommended, or suggested in the labeling of the device. Moreover, FDA is to consider in its determination of effectiveness of the device, whether the extent of the data that otherwise would be required for approval of the PMA, can be reduced through reliance on postmarket controls (section 513(a)(3)(C) of the Act (21 U.S.C. 360c(a)(3)(C)). FDA may rely upon valid scientific evidence, other than a well-controlled investigation, where, for example, such investigation is not reasonably applicable to the device or essential to testing the device (see 21 CFR 860.7(e)(2) and 43 FR 32988 at 32991; July 28, 1978). Congress made clear that it authorized FDA to "accept meaningful data developed under procedures less rigorous than well-controlled investigations in instances in which well-documented case histories assure protection of the public health ..." H.R. Rep. No. 853, 94th Cong., 2d Sess. 1, at 17 (1976). Congress recognized that devices "vary widely in type and in mode of operation, as well as in the scope of testing and experience they have received" and thus, provided latitude to the agency to determine whether the scientific evidence, as a whole, supports safety and effectiveness. Id.

II. ANALYSIS

The SN PMA included data and information on which the agency could conclude there is a reasonable assurance of safety and effectiveness of the BHR system under its conditions of use. SN used a clinical data series to support the safety and effectiveness of the BHR system consisting of 2,385 subjects in which Mr. McMinn implanted the BHR system (the "Overall McMinn Cohort"). The safety assessments included data on revisions, adverse events, deaths, and a metal ion literature review. The effectiveness assessments included survivorship and radiographic data, pain and function data, as evaluated by the Oswestry-modified Harris Hip (OSHIP) Score, and patient satisfaction data. In addition, FDA evaluated unpublished data on 3,374 hips implanted by 140 surgeons and published reports from the experience of multiple surgeons implanting over 3,800 BHR hips around the world.

You raise several issues in your Petition about the sufficiency of the scientific evidence in the SN PMA related to the following topics: 1) single investigator as part of a single investigation; 2) valid scientific evidence; 3) effectiveness measures; 4) least burdensome approach; 5) assessment of design differences in the BHR system; 6) applicability of the data to the U.S. population and U.S. medical practice; 7) consistency of data submitted to FDA for similar devices; 8) financial conflict of interest; and 9) Advisory Panel recommendations. We address your concerns for each below.

A. DATA FROM A SINGLE INVESTIGATOR IN A SINGLE INVESTIGATION

Although you effectively concede that a PMA may be based on data from a single investigator (Petition at 4),[4] you assert that the use of the plural words "investigations," "studies," "trials," "case histories conducted by qualified experts," and "reports" in 21 CFR 860.7(e)(2) means that results from a single investigator do not fulfill the requisite standards for "valid scientific evidence." The terms in the regulation are examples of categories of various types of evidence

[4] You state specifically, that "PMA approvals *generally* require sponsors to present data from more than one investigator in support of a PMA application" (emphasis added), and "If a PMA is based on data from a single investigator ..."

from which it can be fairly and responsibly concluded by qualified experts that there is reasonable assurance of the safety and effectiveness of a device under its conditions of use. The use of plural terms in the regulation does not preclude FDA's reliance on data from one investigation, as you yourself point out in your reference to 21 CFR 814.20(b)(7).[5] Further, you state that "retrospectively-reviewed clinical experience of a single foreign physician" somehow is "unwarranted" and "inconsistent with the governing law" (Petition at 3) and that "other physicians who merely use a device do not constitute investigators whose data can be used in support of a PMA" (Petition at 4). These arguments fail to consider the types of data that support valid scientific evidence. Data on which the agency may rely not only includes data gathered by a physician who participates as an investigator in a well-controlled clinical investigation, but also includes well-documented case histories and reports of significant human experience with a marketed device, whether from a single physician or multiple physicians.

The SN PMA included not only reports on implantations from the Overall McMinn Cohort (with 546 subjects for which there was 5-year follow-up data), but also 3,374 unpublished reports from the Oswestry Outcome Center from 140 surgeons, and published reports from the experience of multiple surgeons implanting over 3,800 BHR systems around the world. The unpublished data on 3,374 hips implanted by 140 surgeons provided additional supportive safety and effectiveness information and helped to demonstrate reproducibility of results. Also, the published literature references which described the use of over 3,800 BHR systems implanted by multiple surgeons in several countries around the world assisted in the development of labeling and helped demonstrate reproducibility of the results.[6] Further, literature information on approved ceramic-on-ceramic THR systems provided a comparison to the BHR system on patient demographics, diagnostic indications, patient accounting, adverse events, revision rates, pain, function, and radiographic results. These data, in several respects, provided the agency with a more complete picture of the clinically significant results and probable benefits compared to probable risks than data that is typically provided to the agency for new or first of a kind orthopedic device, e.g., studies ranging from 100-300 patients (with acceptable clinical and statistical sample size justification) with 2-year follow-up data. Thus, to the extent you argue that evidence from a single investigator and other physicians' experiences with the use of a device do not support an agency finding of valid scientific evidence, such arguments are without merit.

B. VALID SCIENTIFIC EVIDENCE

The agency evaluates the scientific evidence as a whole when it reviews a PMA (21 CFR 860.7(c)(1)). As previously explained, this evidence may vary and must be considered in the context of the particular device and its intended conditions of use. The agency reviewed the SN PMA in the context of the requirements in 21 CFR Part 814. Under 21 CFR 814.15, a study conducted outside the United States submitted in support of a PMA, and not conducted under an

[5] This provision reads in relevant part, "For a PMA supported solely by data from one investigation, a justification showing that the data and other information *from a single investigator* are sufficient ..." (emphasis added). You make a similar argument with respect to 21 CFR 814.15(d)(2) (March 10, 2006 letter at 13), which states "The *studies* have been performed by clinical *investigators* of recognized competence" (emphasis added). Similar to what we say for § 860.20(b)(7), the use of plural terms in the regulation does not preclude our reliance on data from one investigator.

[6] The unpublished and published data on the BHR system are not merely "anecdotal evidence" (Petition at 3 and 20), but rather reports of significant human experience with a marketed device, consistent with 21 CFR 860.7(c)(2).

Investigational Device Exemption (IDE), shall comply with the applicable requirements in such section.[7] Whether or not one characterizes the evidence submitted by SN as a "well-controlled investigation" or as a combination of any of the descriptors listed in 21 CFR 860.7,[8] FDA concluded that the data as a whole provides valid scientific evidence that is adequate to support a reasonable assurance that the device is safe and effective for its intended conditions of use (21 CFR 860.7(c)(2)).

This device is intended for patients requiring primary hip resurfacing arthroplasty, due to: 1) non-inflammatory arthritis (degenerative joint disease) such as osteoarthritis, traumatic arthritis, avascular necrosis, or dysplasia/DDH; or 2) inflammatory arthritis such as rheumatoid arthritis. The BHR system is intended for patients who, due to their relatively younger age or increased activity level, may not be suitable for traditional total hip arthroplasty due to an increased possibility of requiring future ipsilateral hip joint revision. We address, below, several of the concerns you raise related to patient selection, reproducibility, and potential bias, based on the data submitted in the SN PMA.

The sufficiency of the scientific evidence is outlined in the Summary of Safety and Effectiveness (SSED) (Appendix 1). The application included sufficient information to establish an appropriate patient population for the intended use and conditions of use of the BHR system (see SSED at 9-10). To develop the indications for use and physician labeling from experience gained from the 2,385 implantations, SN provided a list of the factors that contributed to the investigator's decision to perform a total hip replacement (THR) rather than the BHR hip resurfacing procedure (SSED at 9-10). The investigator collected a patient's pre-operative history, physical, and diagnostic work-up to screen candidates for BHR versus THR procedure. FDA considers the list of factors, when considered in the context of other controls related to investigator's collection of other data and information related to these procedures and other literature, sufficient on which to base labeling, including warnings and precautions.[9]

The information provided was on a consecutive case series of 2,385 patients (the Overall McMinn Cohort) with sufficient information on a large number of subjects and length of follow-up, for which FDA could evaluate the primary effectiveness endpoint of survivorship, i.e., time-

---

[7] The agency rejects your argument that suggests SN data had to comply with the Investigational Device Exemption regulations in 21 CFR Part 812. FDA will accept studies submitted in support of a PMA conducted outside the United States and not conducted under an IDE, provided that the studies are conducted in accordance with the "Declaration of Helsinki, " as incorporated into 21 CFR 814.15, or the laws and regulations of the country in which the research are conducted (21 CFR 814.15). The laws and regulations of the UK (including the Data Protection Acts of 1984 and 1998) were followed.

[8] Section 860.7 lists valid scientific evidence as including evidence from well-controlled investigations, partially controlled studies, studies and objective trials without matched controls, well-documented case histories conducted by qualified experts, and reports of significant human experience with a marketed device.

[9] Your petition stated that "it is unlikely that, in absence of a protocol and predefined forms, records collected by different physicians were recorded and presented in a consistent format or with identical types of information included." The investigator's evaluations, however, were conducted according to standard medical practice, and the investigator performed all but one revision surgery. Other data collected included prospectively collected adverse event data in annual patient questionnaires. These data, along with still other data, allowed for the collection and presentation of the safety data (i.e., revisions and adverse events) over time (SSED at 22-24). The presentation of these data over time provides clinical context to this information (contrary to your assertion in Petition at 13, note 68).

specific cumulative revision-free probability.[10] The data, which included the 2,385 implantations, the factors considered for when to implant a BHR or THR, and an analysis of the BHR revisions in the 2,385 implantations (i.e., femoral neck fracture, femoral head collapse, dislocation, AVN, and infection), were used, in part, to develop the indications for use and physician labeling. In addition, literature findings were also considered (SSED at 10) concerning certain device-related failures, upon which SN included warnings and precautions in the labeling.

In addition, SN compared their results for the BHR system to two literature controls for ceramic-on-ceramic total hip systems (i.e., the literature articles contained information on some of the cases presented in the Howmedica Osteonics ABC System and Trident System PMA, P000013 and Wright Medical's Ceramic Transcend Articulation System PMA, P030027) (see SSED at 17). Although there were some differences between the data in the literature concerning the approved ceramic-on-ceramic THR systems and the data for the BHR clinical data series, these differences did not preclude FDA's reliance on such literature to provide additional valuable information to assist the agency in generalizing the findings on the BHR system to the appropriate patient population for which the BHR system is intended, particularly with respect to the findings in the literature related to patient demographics, diagnostic indications, patient accounting, adverse events, revision rates, pain, function, and radiographic results. The literature data served as historical controls from which the agency could compare data from patients who received the BHR system (contrary to what you assert in Petition at 20).[11]

The data provided from the Overall McMinn Cohort, along with literature reviewed for labeling, including warnings and precautions, and the comparison of the patient population to the U.S. population in the Overall McMinn Cohort supported FDA's conclusion that the data was sufficient to ensure the reproducibility of test results. Contrary to what you assert (Petition at 20), FDA did not rely solely on the data from the 3,374 unpublished reports from the Oswestry Outcome Center from 140 surgeons to ensure reproducibility. In addition to data from the Overall McMinn Cohort and these unpublished reports, FDA relied on published reports from the experience of multiple surgeons implanting over 3,800 BHR systems around the world, which provided supportive safety and effectiveness information and helped to demonstrate reproducibility of the results. The data, as a whole, provide adequate assurance of reproducibility of test results. FDA did not rely on "mere anecdotal report of other physicians' use" to demonstrate reproducibility (Petition at 20). Moreover, the agency is requiring certain post-market studies to evaluate the learning curve and training program for U.S. surgeons and to evaluate the longer-term safety and effectiveness.

The racial and ethnic distributions in the U.S. and U.K. populations are similar. There were noted differences in the higher percentage of people of African-descent and "other races" in the general U.S. population as compared to the general U.K. population. However, any differences were not

---

[10] The following references provide additional information regarding the evaluation of survivorship:
Lee, Elisa: Statistical Methods for Survival Data Analysis, Wiley Interscience 1992
Klein, John P. and Moeschberger, M.L. Survival Analysis: Techniques for Censored and Truncated Data. Springer 1997.

[11] Specifically, you seem to suggest that patients receiving the BHR devices should have been compared to those who received the conventional hip prostheses by Mr. McMinn, and that therefore, there was no control group from which labeling could be written. However, you fail to consider other types of controls that may be adequate.

significant when considering other supportive unpublished data provided on 3,374 BHR hips implanted by 140 surgeons and published reports from the experience of multiple surgeons implanting over 3,800 BHR hips in several countries around the world. The data and information as a whole supports the applicability of data to the U.S. population and medical practice. In addition, the practice of medicine, specifically the orthopedic practice of medicine, utilized by the investigator is considered to be similar to the standard of orthopedic practice in the U.S. The applicability of the foreign data to the U.S. patient population is based on its large sample size, as well as the comparable demographics and diagnostic indications to the multi-center literature reference group. Thus, contrary to your assertion that "reproducibility cannot be demonstrated by mere use of a device by more than one physician" (Petition at 8), the data as a whole are sufficient to ensure reproducibility. Although reproducibility is supported by the data and information in the SN PMA, FDA is requiring additional post-approval study of SN's training program, which information must be submitted for evaluation by FDA every 6 months for two-years.

Lastly, SN identified several steps that it took to minimize potential bias in the study, including:[12]

- The PMA is based on a combination of prospective and retrospective data collection of the complete patient series implanted within pre-specified time frames. No patient implanted with the BHR at the McMinn Centre was excluded from data summary, presentations, and analysis;
- The clinical evaluations at follow-up intervals (OSHIP questionnaire) were conducted by the Oswestry Outcomes Center, an independent outcomes research organization, and not the implanting physician (Mr. McMinn);
- The clinical evaluations are based on patient self-assessment instead of physician-administered assessment;
- The radiographic evaluations were conducted centrally by an independent radiologist and not by the implanting physician or another physician at the same clinic (The McMinn Centre). Further, the independent radiologist practices at a different hospital than the study investigator (Mr. McMinn); and
- Data were audited by S&N and FDA.

Based on the reasons set forth above, these data and information that the agency relied upon in the SN PMA, when combined with the other data and information in the application and available to FDA, were sufficient to support the agency's finding of safety and effectiveness.

[12] The reference to General Medical Co. v. FDA, 770 F.2d 214 (D.C. Cir. 1985) that you cite to for the proposition that a lack of a protocol, including insufficient patient follow-up, may indicate a potentially biased study is not relevant to the data under consideration in the SN PMA. The follow-up rate for the SN PMA were high enough to demonstrate that the information was representative of the target population (i.e., 90.8 % follow-up was achieved for the first 1626 patients in the Combined X-Ray / Oswestry Cohort, upon which the effectiveness analyses were performed; this group of patients included 546 procedures (hips) evaluated at 5 years in this cohort).

C. MEASURES OF EFFECTIVENESS

FDA determined that the information provided by SN to address pain and function, in conjunction with the 5-year survivorship, radiographic, and patient satisfaction data, are reliable and accurate and provide a reasonable assurance of the effectiveness of the device.
SN used an instrument, the Oswestry-modified Harris Hip (OSHIP) score, in which patients provide a self–assessment of pain and function. SN provided information to address how the 5-year OSHIP data was collected, how the OSHIP scoring system was developed, and provided published and unpublished information to justify its use (see SSED at 14-16). After comprehensive review of the information provided, including published and unpublished evaluations of this instrument, FDA concluded that this instrument is an acceptable measure of patient outcome. Data are available on over 90% of the patients that were used to evaluate pain and function with the OSHIP scoring system. Literature are available indicating that the OSHIP scores tend to be lower when compared to the use of the Harris Hip Score; reflecting the likelihood that OSHIP would provide, if anything, a more conservative measure of pain and function. Id. In reviewing these data, in conjunction with the 5-year survivorship, radiographic, and patient satisfaction data, FDA determined that the OSHIP scoring system is a reliable and accurate measure of pain and function which the agency could use to assess the efficacy of this device.

The survivorship estimates were based on the number of patients with no revision. Survivorship statistical analyses were provided for various cohorts (i.e., X-ray (n=124), Oswestry (n=1502), McMinn (n=759), Combined X-ray/Oswestry (n=1626), Overall McMinn (n=2385)) and demographic subgroups (i.e., subgroups based on gender, age, diagnostic indication, unilateral/bilateral, baseline OSHIP score, body mass index (BMI)). There were no statistically significant differences in cumulative 5-year survival (revision-free) probabilities among three study cohorts (i.e., X-ray, Oswestry, McMinn). Survivorship for the Combined X-ray/Oswestry and Overall McMinn Cohorts were 98.4% (95% confidence interval=97.3-99.5%) and 98.5% (95% confidence interval= 97.4-99.6%), respectively. The only marginally statistically significant difference in 5-year survival probability was between the patients with Osteoarthritis (98.8%) and Avascular Necrosis (92.1%) as their primary diagnostic indication. In this case series, survivorship results on a large number of patients at 5-years were comparable to most 2-year results for other hip studies. In addition, survivorship helped in an analysis of which types of patient groups benefit the most from a resurfacing device.

In addition to the measure of survivorship, other efficacy endpoints for the case series included measures of pain and function (discussed in section II.C), radiographic outcome, and patient satisfaction. Regarding the radiographic data, a radiographic study protocol was developed prior to evaluation of the radiographs and included a prospective definition for radiographic success/failure. Your characterization of these data as taken at "non-standardized intervals and evaluated according to post-hoc criteria" is unfounded (Petition at 18). The data were evaluated after developing a protocol for the retrospective analysis. Moreover, the agency determined the evaluation of the radiographic data in the early post-operative period and again at 5-years post-operation to be acceptable, given the longer 5 year follow-up. The PMA contained the results of an independent radiographic review of the X-Ray Cohort, the first 124 procedures performed in the series from July 1997 through December 1997. Radiographs were taken on 108 of the 118

procedures expected at 5 years postoperatively (91.5%). Baseline films for the purposes of comparison were made in each of the 108 cases in the postoperative time period, usually within 3 months.[13] The X-Ray Cohort of 124 hips included 81 men (65.3%), 43 women (34.7%), with diagnostic indications of OA (n=92 (74.2%)); dysplasia/developmental dislocation of the hip (DDH) (n=22 (17.7%)); avascular necrosis (AVN) (n=7 (5.6%)); inflammatory arthritis (n=2 (1.6%)); and other (n=1 (0.8%)). The rates of these diagnoses are comparable to the rates in the larger BHR cohorts, and thus reflect the proposed indications for the use of the BHR system. These data were more extensive than what FDA typically relies on to approve new or first of a kind orthopedic devices with studies, which range from 100 – 300 patients (with acceptable clinical and statistical sample size justification) with only 2-year follow-up data, in that they evaluated radiographic outcomes at 5 years. Therefore, your assertion (Petition at 17-18) that the "radiographic data presented do not include evidence of the effectiveness of the BHR in patients with other forms of non-inflammatory arthritis... or with inflammatory arthritis" is not supported.

Patient satisfaction data were also collected using the annual, patient-completed, mail-in questionnaire. For the purpose of the BHR study, an additional question about patient satisfaction was appended to the end of the OSHIP assessment questionnaire. At 5 years, 99.5% of the procedures in the X-Ray/Oswestry combined cohort were pleased or very pleased with the operation.

D. THE "LEAST BURDENSOME" APPROACH

You state in your petition that the "'least burdensome provision'[14] notwithstanding, a class III medical device can be approved through the PMA process only if the sponsor demonstrates, through valid scientific evidence, the safety and effectiveness of the device" (Petition at 11). You also stated that the least burdensome approach for this device can only be a well-controlled clinical trial. Id. at 11. We agree with your assertion that a class III medical device can be approved only when there is valid scientific evidence to demonstrate that the device has a reasonable assurance of safety and effectiveness. As discussed above, the valid scientific evidence submitted for SN's PMA does demonstrate this. However, your discussion in the Petition about why you believe a well-controlled clinical trial is the "least burdensome" approach is not relevant to whether the data and information submitted by SN in the PMA is actually valid scientific evidence that demonstrates a reasonable assurance of safety and effectiveness. As explained earlier (Section II.B), FDA reviewed the data as a whole and determined that the data provides valid scientific evidence that is adequate to support a reasonable assurance that the device is safe and effective for its conditions of use.

E. THE DESIGN DIFFERENCES IN THE BHR SYSTEM

As with any PMA, in order to demonstrate that there is reasonable assurance that a device is effective, clinically significant results are to be provided in a significant portion of the target

[13] Although we would have preferred immediate post-operative films to be used as the baseline x-ray data, the films taken in the early post-operative period were deemed sufficient to use as baseline for comparison to the 5-year films.
[14] Section 513(a)(3)(D)(ii) states that FDA must consider "the least burdensome appropriate means of evaluating device effectiveness that would have a reasonable likelihood of resulting in approval." (21 U.S.C. 360c(a)(3)(D)(ii).)

population (21 CFR 860.7 (e)(1)). SN provided adequate justification for including the variations in the BHR system in its labeling for the device.[15] The BHR prosthesis is a metal-on-metal hip resurfacing device that consists of a stemmed femoral head resurfacing component designed for cemented fixation, and a hemispherical acetabular cup designed for cementless, press-fit, fixation (SSED at 2). The resurfacing head, acetabular cups and dysplasia cup screws come in various sizes (SSED at 3). Although the data from the clinical series were not stratified based on acetabular cup type (i.e., standard, dysplasia, and bridging acetabular components), it was not necessary to do so in order to determine safety and effectiveness of the various acetabular cup types because the dysplasia and bridging cups (which were implanted in 174 of the 2,385 hips) are primarily designed for patients with deficiency in the superolateral aspect of the acetabulum (i.e., patients with dysplasia/developmental dislocation of the hip (DDH)) and survivorship, pain and function, and patient satisfaction data were stratified for these patients. This information is summarized in the labeling for the device (SSED at 31 for survivorship, at 36 for pain and function, and at 38 for patient satisfaction and is also reproduced in the physician labeling) and provides the physician with important information for patients with dysplasia/DDH and who may receive the dysplasia or bridging acetabular cup designs.

F. APPLICABILITY TO U.S. POPULATION AND U.S. MEDICAL PRACTICE

FDA may approve a PMA based solely on foreign clinical data, among other criteria, if the foreign data are applicable to the U.S. population and U.S. medical practice (21 CFR 814.15). As previously discussed (Section II. B), the data are applicable to the U.S. population. Although the BHR system has not been used in the U.S. before, the practice of medicine, specifically the orthopedic practice of medicine, utilized by the investigator is considered to be similar to the standard of orthopedic practice in the U.S. A summary of the investigator's practice of medicine was provided for reference in the SSED (SSED at 39).

In making a determination of a reasonable assurance of the effectiveness of a device, FDA is to consider whether the extent of data that otherwise would be required for approval can be reduced through reliance on post-market controls (section 513(a)(3)(C) of the Act (21 U.S.C. 360c(a)(3)(C)). FDA may impose postapproval requirements in a PMA approval order at the time of approval of the PMA (21 CFR 814.82).

FDA is requiring several post-market controls that include the conduct of a study to evaluate the learning curve, training program, and longer-term data on the BHR system in the United States. This study is expected to include 350 patients at up to 8 sites with a minimum of 35 patients per site. The study will include 5-year annual post operative follow-up, interim follow-up in years 6-9, and at again at 10 years. As part of the training program, investigators will be recruited from a "Core Surgeon" group who will be initially trained on the BHR system and other interested U.S. surgeons trained from this core group. As part of the training program, quarterly teleconferences or meetings among investigators are to occur for the first two years, in part, to identify recommendations for improvement of the training program or the labeling.

[15] FDA did not hold Wright to a different standard, as you assert in the October 11, 2005 letter at 5. FDA asked that you provide clinical support for the indications or adequate justification for including the indications in the labeling. SN provided adequate justification for including its indications in its labeling.

G. CONSISTENCY OF DATA IN PMAs FOR TOTAL HIP RESURFACING DEVICES

Your petition also asserts that approval of the SN PMA would be inconsistent with FDA precedents for class III orthopedic devices. In support of your assertion, you refer to the preamble of a 2004 proposed rule for hip joint metal/polymer or ceramic/polymer semi-constrained resurfacing cemented prosthesis (69 FR 10390, March 5, 2004). In that preamble, we stated that a PMA should include valid scientific evidence as defined in 21 CFR 860.7, obtained from well-controlled studies or another form of valid scientific evidence (69 FR at 10394). We also provided recommendations related to general clinical study considerations. Thus, to the extent that preamble is relevant to SN's PMA for a metal-on-metal hip resurfacing device, it also supports approval of the PMA as long as the PMA is supported by valid scientific evidence. As discussed above, we have determined that there is valid scientific evidence to support SN's PMA.

Your October 11, 2005, letter alludes to the IDE and data we requested for the Wright Medical CONSERVE® Plus PMA. Consistent with the Act, we informed you that an IDE was required. Studies of investigational devices in this country are subject to 21 CFR parts 50, 56, and 812 (see also section 520(g) of the Act; 21 U.S.C. 360j(g)). As discussed earlier, studies conducted outside the U.S. and not conducted under an IDE, are subject to the provisions of 21 CFR 814.15, which requires that the studies be conducted in accordance with the "Declaration of Helsinki" or the laws and regulations of the country in which the research are conducted. The laws and regulations of the U.K. (including the Data Protection Acts of 1984 and 1998) were followed.

If FDA determines the data and information in a PMA support safety and effectiveness, under relevant statutory and regulatory provisions where none of the grounds for denying approval exist, FDA must issue an order approving the application (section 515(d)(1)(A) of the Act (21 U.S.C. 360e(d)(1)(A)). FDA's reliance on data of the type submitted in SN's PMA is not inconsistent with its past practice for other Class III device approvals.[16] Thus, FDA's determination that the data and information in the SN PMA is valid scientific evidence is not "unprecedented" (Petition at 3).

FDA considered the data and information in the SN PMA as a whole. Based on its review, FDA determined that there is a reasonable assurance of safety and effectiveness for the BHR System under the conditions of use prescribed, recommended or suggested in the proposed labeling.

---

[16] P000057, Ascension Orthopedics, Inc., Ascension MCP (semi-constrained finger); P960054, Johnson & Johnson Professional, Inc., S-ROM Poly-Dial Constrained Liner; and P960053, Avanta Orthopaedics, Braun-Cutter Trapezo-Metacarpal Prosthesis.

H. FINANCIAL DISCLOSURE OF THE INVESTIGATOR

Financial disclosure information was provided in the PMA in accordance with 21 CFR 814.20(b)(12). SN identified the steps taken to minimize potential bias in the study (discussed in section II.B).

In addition, FDA performed an inspection of Mr. McMinn's clinical site (McMinn Center) and the Oswestry Outcomes Center. FDA did not find any reason to believe that there was anything other than valid and reliable data for the purposes of evaluating the safety and effectiveness of the BHR System.

III. ADVISORY PANEL RECOMMENDATIONS

FDA fully considered the Advisory Panel evaluation and recommendations, both pro and con, in the agency's decision to approve the SN PMA. The Orthopaedic and Rehabilitation Devices Panel (the Panel) met on September 8, 2005 in Gaithersburg, MD to make a recommendation to the FDA on the approvability of the Smith and Nephew, Inc. BHR system, P040033. FDA received expert clinical opinion from the Panel regarding the safety and effectiveness data collection methods, the applicability of the foreign data from a single investigator and United Kingdom practice of medicine to the target United States population and practice of medicine, and the study results with respect to the device's safety and effectiveness.

Regarding the safety data collection methods, some Panel members expressed concerns about what they considered a lack of prospectively collected information and that the safety information was coming from one source. The applicant clarified that the Oswestry Outcome Center collected data prospectively starting in 1997 and an independent group compiled all safety data for the case series retrospectively. Other Panel members indicated that because of the large series and the corresponding amount of safety data included, there was enough information to evaluate device safety. Also, Panel members stated that the methods used met the FDA definition of valid scientific evidence in 21 CFR 860.7.

Regarding the effectiveness data collection methods, some Panel members supported the use of patient questionnaires to capture safety and effectiveness data stating that it may reduce bias associated with patients wanting to please their physicians in their responses to physician administered questionnaires. Other Panel members stated that although patient questionnaires are important, physical exams and radiographic data is also as important; these Panel members would have used a more conventional evaluation (i.e., HHS) rather than OSHIP.

Regarding whether or not the clinical data is applicable to the target U.S. population, practice of medicine, and U.S. orthopedic surgeon population, some Panel members expressed concerns that the case series did not contain data on the variability of the use of the device at various centers or that the data would provide reassurance of its applicability to the U.S. population at risk. Other Panel members expressed support for the clinical data stating: (1) that if a randomized, controlled, trial would have been performed in the U.S. for this device, the applicant would have recruited a small group of five excellent hip surgeons who would be very different from the average hip surgeon; therefore, the only difference is that data has one learning curve rather than

five; (2) that the U.K. population appeared very similar to the U.S. population including the use of referral practices; (3) that there was additional literature information on the use of the device by other surgeons; and (4) that an adequate training plan and proposed post-approval study could mitigate these concerns.

The Panel voted three to two to recommend that FDA approve the PMA with conditions. The recommended condition of approval was as follows: The applicant should conduct the proposed post-approval study presented in the PMA with the addition of a clinical and radiographic evaluation at the 10-year follow-up time point. In addition, the sample size for the post-approval study should be based on statistical principles and the criteria for success. FDA's approval order has incorporated these Panel recommendations in its conditions of approval. Information concerning this Panel meeting, including a transcript can be found at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfAdvisory/details.cfm?mtg=600.

IV. CONCLUSION

FDA has reviewed your petition, the submitted comment, and other relevant data and information available to the agency. For the reasons discussed above, your request that we deny approval of SN's PMA for the BHR system is denied.[17]

If you have any questions about this response, please contact Myrna Hanna of our Regulations Staff at (240) 276-2347.

Sincerely yours,

Linda S. Kahan

Linda S. Kahan
Deputy Director
Center for Devices and
Radiological Health

Enclosure

[17] SN raised an issue in their February 8, 2006 letter about the use of the review process in section 515(g) of the Act (21 U.S.C. 360(e)(g)) as a means for addressing concerns about the scientific evidence submitted in a PMA, rather than the use of the citizen petition process under 21 CFR 10.30. This response to your citizen petition should not be interpreted as a determination that the agency could not have relied on the 515(g) review process.

Summary of Safety and Effectiveness Data

I. GENERAL INFORMATION

Device Generic Name: Prosthesis, Hip, Semi-constrained, Resurfacing, metal/metal, hybrid fixation

Device Trade Name: Birmingham Hip Resurfacing (BHR) System

Applicant Name and Address: Smith & Nephew, Inc. Orthopaedics Division
1450 Brooks Road
Memphis, Tennessee 38116

PMA Number: P040033

Date of Panel Recommendation: September 8, 2005

Date of Notice of Approval to Applicant: May 9, 2006

II. INDICATION FOR USE

The Birmingham Hip Resurfacing (BHR) System is a single use device intended for hybrid fixation: cemented femoral head component and cementless acetabular component. The BHR system is intended for use in patients requiring primary hip resurfacing arthroplasty due to:

- Non-inflammatory arthritis (degenerative joint disease) such as osteoarthritis, traumatic arthritis, avascular necrosis, or dysplasia/developmental dislocation of the hip (DDH), or
- Inflammatory arthritis such as rheumatoid arthritis.

The BHR System is intended for patients who, due to their relatively younger age or increased activity level, may not be suitable for traditional total hip arthroplasty due to an increased possibility of requiring future ipsilateral hip joint revision.

III. CONTRAINDICATIONS

- Patients with infection or sepsis
- Patients who are skeletally immature
- Patients with any vascular insufficiency, muscular atrophy, or neuromuscular disease severe enough to compromise implant stability or postoperative recovery
- Patients with bone stock inadequate to support the device including:
  - Patients with severe osteopenia should not receive a BHR procedure. Patients with a family history of severe osteoporosis or severe osteopenia.

- Patients with osteonecrosis or avascular necrosis (AVN) with >50% involvement of the femoral head (regardless of FICAT Grade) should not receive a BHR.
- Patients with multiple cysts of the femoral head (>1cm) should not receive a BHR.
- Note: In cases of questionable bone stock, a dual-energy x-ray absorptiometry (DEXA) scan may be necessary to assess inadequate bone stock.

- Females of child-bearing age due to unknown effect on the fetus of metal ion release
- Patients with known moderate to severe renal insufficiency
- Patients who are immunosuppressed with diseases such as AIDS or persons receiving high doses of corticosteroids
- Patients who are severely overweight
- Patients with known or suspected metal sensitivity (e.g., jewelry)

IV. WARNINGS AND PRECAUTIONS

- Patients on medications (such as high-dose or chronic aminoglycoside treatment) or with co-morbidities (such as diabetes) that increase the risk of future, significant renal impairment should be advised of the possibility of increase in systemic metal ion concentration. Preoperative and postoperative monitoring of renal function (such as creatinine, glomerular filtration rate (GFR), blood urea nitrogen (BUN)) will be necessary.

- Only physicians who have received appropriate training and are familiar with the implant components, instruments, procedure, clinical applications, adverse events, and risks associated with the BHR System should use this device. Contact Smith & Nephew, Inc. for the surgical technique manual and procedural training protocol.

- Currently, Smith & Nephew, Inc. does not have a commercially available modular metal femoral head for use with a BHR resurfacing shell. Therefore, if the BHR resurfacing head must be revised to a total hip arthroplasty, the acetabular shell should also be revised, even if well fixed.

See complete list of *Warnings* and *Precautions* in the Instructions for Use for the BHR System.

V. DEVICE DESCRIPTION

The Birmingham Hip Resurfacing (BHR) prosthesis is a metal-on-metal hip resurfacing prosthesis. The device consists of a stemmed femoral head resurfacing component designed for cemented fixation, and a hemispherical acetabular cup designed for cementless, press-fit, fixation. Both components are manufactured from high carbon, as-cast, cobalt chrome (CoCr) alloy (ASTM F75, *Specification for Cobalt-Chromium-Molybdenum Alloy for Surgical Implant Applications,* and ISO 5832-4, *Implants for Surgery – Metallic Materials – Part 4*). The BHR is a "resurfacing" prosthesis because only the surface of the femoral head is removed to implant the femoral head resurfacing component. The acetabular cups are configured in standard, dysplasia, and bridging designs. All acetabular cups have a single layer of integrally-cast CoCr-alloy (ASTM F75 and ISO 5832-4) beads on the outer surface that are coated with hydroxyapatite (HA) (ASTM F1185, *Standard Specification for Composition of Hydroxylapatite for Surgical Implants*). Instrumentation sets are provided as standard; several additional instruments are available as options.

Resurfacing Femoral Head

The resurfacing femoral head is supplied in a range of six sizes. The femoral head central stem is parametric and varies proportionally with the external diameter. There are 6 equally spaced internal recesses intended to provide antirotational locking for the cement mantle.

Acetabular Cups

The standard acetabular component is supplied in a range of twelve sizes (two for each femoral head size to address the condition of occasional head cup mismatch). For those patients with a deficiency in the superolateral aspect of the acetabulum, the dysplasia cup is available. The dysplasia cup is designed with two superolateral screw holes that accommodate CoCr-alloy dysplasia cup screws. There is a range of six sizes for the dysplasia cup. A bridging cup is designed with a thicker wall section than the dysplasia cup to allow for mismatch between femoral head size and surgically prepared acetabulum. The bridging cup is also designed with two superolateral screw holes that accommodate the CoCr-alloy dysplasia cup screws. The bridging cup is available in five sizes.

Dysplasia Cup Screws

The dysplasia cup screws are threaded through a threaded lug on the superolateral aspect of either the dysplasia or bridging cup and lock in situ. The screws also lock into the posterior cortical bone of the ilium. Screws are available in sizes ranging from 24mm to 88mm, in 2mm increments.

Sizing and System Compatibility

Each femoral head resurfacing component is compatible with two standard acetabular cup sizes and one dysplasia or bridging cup size as shown in Table 1 below.

Table 1: BHR Head and Cup Sizing and System Compatibility

BHR Femoral Head Resurfacing Component (identified by head outer diameter)	Mating BHR Standard Cup Sizes (2 cups available per head component size)	Mating BHR Dysplasia Cup Sizes	Mating BHR Bridging Cup Sizes
38mm	44mm or 46mm	46mm	50mm
42mm	48mm or 50mm	50mm	54mm
46mm	52mm or 54mm	54mm	58mm
50mm	56mm or 58mm	58mm	62mm
54mm	60mm or 62mm	62mm	66mm
58mm	64mm or 66mm	66mm	

VI. ALTERNATIVE PRACTICES OR PROCEDURES

Depending on individual circumstances, alternative procedures may include non-surgical treatment such as reduced activity and/or pain management; other surgical treatments that do not involve use of an implant such as a Girdlestone procedure; or use of other commercially available total hip replacement systems. Commonly used implant bearing materials for total hip arthroplasty include metal on ultra-high molecular weight polyethylene (UHMWPE), ceramic on

UHMWPE, metal on metal, and ceramic on ceramic. Total hip prostheses are implanted by either cemented or uncemented techniques.

VII. MARKETING HISTORY

The BHR is marketed in Australia, Austria, Belgium, Bermuda, Canada, Denmark, Egypt, Germany, Hong Kong, India, Ireland, Israel, Italy, Japan, Korea, Netherlands, Portugal, Romania, South Africa, Spain, Sweden, and Switzerland. The BHR has not been withdrawn from any country due to safety and effectiveness reasons.

VIII. POTENTIAL ADVERSE EFFECTS OF THE DEVICE ON HEALTH

Reported Device Related Adverse Effects

The most commonly reported BHR device related adverse events are:

- femoral neck fracture
- femoral head collapse
- infection
- avascular necrosis
- dislocation
- component migration/loosening, and
- impingement

A complete list of the frequency and rate of complications and adverse events identified in the case series review is provided below in Section X: Summary of Clinical Studies, Tables 17 and 18.

Potential Adverse Effects

The following adverse effects may occur in association with hip replacement surgery including the BHR System:

- Cardiovascular complications including venous thrombosis, pulmonary embolism, or myocardial infarction
- Sudden, pronounced, intraoperative blood pressure decrease due to the use of bone cement,
- Hematoma or damage to blood vessels resulting in large blood loss
- Delayed wound healing
- Superficial or deep infection. Infections may occur months to years after surgery and these infections are difficult to treat and may require reoperation with removal surgery and later replacement at another time
- Temporary or permanent nerve damage resulting in functional and/or sensory deficits in the affected limb
- Metal sensitivity reactions or allergic reactions or metallosis
- Dislocation or subluxation leading to post-operative joint instability (which may be caused by malpositioning of the implants, or muscle or fibrous tissue laxity)

- Component loosening or migration due to trauma, loss of fixation, malalignment, or bone resorption
- Limb length discrepancy
- Increased hip pain and/or reduced hip function
- Fatigue fracture of the implants as a result of excessive loading, malalignment, or trauma,
- Osteolysis and/or other peri-prosthetic bone loss
- Unintended bone perforation or fracture occurring either intra-operatively or post-operatively as a result of trauma, excessive loading, osteolysis, or osteoporosis
- Periarticular calcification or ossification
- Wear or deformation of the articular surface as a result of excessive loading or implant malalignment

Any of these adverse effects may require medical or surgical intervention. Rarely, these adverse effects may lead to death.

IX. SUMMARY OF PRECLINICAL STUDIES

Non-clinical laboratory information was provided in support of the BHR System, including information on biocompatibility, wear, friction, surface topography, kinematics, component stress analysis, beaded surface and HA coating characterization, explant analysis, sterilization and shelf-life.

Biocompatibility Studies

The chemical composition of the BHR device is defined in ISO 5832-4, *Implants for Surgery – Metallic Materials – Part 4*, and ASTM F75, *Specification for Cobalt-Chromium-Molybdenum Alloy for Surgical Implant Applications*, and is outlined in Table 2 below.

Table 2: BHR Component Chemical Composition

Chromium 27-30%	Molybdenum 5-7%	Nickel 0.5% max.	Iron 0.75% max.
Carbon 0.35% max.	Manganese 1% max.	Silicon 1% max.	Cobalt balance

Technical specifications for hydroxyapatite coating of the BHR acetabular cup meet the requirements for ISO 13779-1, ISO 13779-2, and ISO 13779-4 (draft), *Implants for Surgery - Hydroxyapatite - Parts 1, 2 and 4*. Because the device is comprised of well-accepted materials for a permanent implant, and meets the ISO standard, additional biocompatibility testing is not required.

Articulating Wear, Friction, and Surface Topography Analyses

A tribological study of the BHR was conducted to analyze volumetric wear rates for 5 million cycles. The study used a hip joint simulator to compare the volumetric wear rates of five devices subjected to dynamic loads and motions and one control specimen that was dynamically loaded but experienced no tangential motions. Also, with a hip friction simulator, friction was measured before, during and after the wear test. In addition, surface topography of the bearing surfaces before and throughout the wear simulator study was evaluated.

All femoral head prostheses were 50mm outer diameter size components. Taking into account competing factors such as sliding distance, fluid-film lubrication thickness, fluid entraining velocity, and clearance between the head and acetabular cup, the 50mm components were selected as representative of a "worst-case" wear couple. The joints were tested in the anatomical position (the femoral component being below the acetabular component) with the acetabular component oriented at 33° to the horizontal. The force vector applied provided a minimum load of 100 N and the maximum load was 2975 N.

Over the 5 million cycles, the average wear rate was 1.33 mm^3/million cycles. Initially the wear rate was high compared with later cycles (0.5 to 1 million cycles, 3.00 mm^3 /million cycles), but by 3-5 million cycles the wear rate had reduced to 0.4mm^3/million cycles. The results were compared to other total metal on metal hip systems. The wear testing, in conjunction with the clinical data and metal ion evaluations summarized in the "Summary of Clinical Data" section below, provided an acceptable characterization of the wear performance of the device.

On one of the head/cup couples, friction tests were carried out before, during and after wear testing. A modified Paul curve was used to provide a dynamic loading cycle with a maximum load of 2000N and a minimum load of 100N. The femoral head was in an inverted position but with a relative position between the head and cup the same as the wear simulator. As flexion-extension motion took place, the friction generated within the prosthesis was measured throughout the cycle. Tests were performed using several lubricants with varying degrees of viscosity. Decreasing friction over the course of the test was reported. Frictional torque ranged from 4.48-4.81Nm pre-test, to 0.75-1.88Nm after 3 million cycles, to 0.89-1.32Nm after 5 million cycles. Frictional torque appeared to be a bit higher for lower lubricant viscosities but this was not consistent for all components tested.

Surface topography was measured by profilometry before and throughout the wear simulator study. As supplied, the heads exhibited surface topography with irregular shaped peaks often seen as rectilinear arrays that looked like carbides within the CoCrMo. The peak-to-valley heights (PV) averaged 0.320 (SD 0.081) μm with a positive skewness, indicating that the majority of the height ranges were above the mean line and thus were peaks. It was indicated that these surfaces are typical for this alloy system where the surface has "relief polishing" due to the differences in hardness between the carbide and matrix phases. At the end of the test the PV values had generally increased for the heads. In the case of the cups, some increased and some decreased. The positively skewed distribution for the heads had generally worn to a negative one by the end of the test, indicating that most variations from the mean plane were in the form of scratches while the peaks had been smoothed. The skew had decreased for all the cups also but still remained slightly positive.

In summary, the wear rates, frictional results and surface topography of the BHR hip resurfacing device are within the range of other metal-on-metal total hip replacement (THR) bearings cleared through premarket notification and results reported in published literature.

Kinematics

The range of motion (ROM) test procedure was performed according to that described in Annex A of ISO 21535: 2002 (EN 12563: 1998), *Non-Active Surgical Implants - Joint Replacement Implants - Specific Requirements for Hip-Joint Replacement Implants*. ISO 21535 was written for diaphyseal anchored types of hip implants; therefore, the test procedure was modified because the femoral implant component is attached to the proximal femoral head directly and not to an intramedullary stem. Since the ROM of the resurfacing head is restricted by impingement

between the femoral neck and the rim of the acetabular cup, to simulate the femoral neck, a cylindrical plastic component with the inner diameter of the femoral head was placed on the stem of the femoral head. The ROM is limited by the contact between the cylindrical "femoral neck" and the rim of the acetabular cup. A rationale for components and methods as representative of "worst case" scenario determined the 58mm femoral head size paired with the 64 or 66mm acetabular cup to have the smallest angular displacement. For the purpose of this test the 66mm cup was utilized.

The test result for flexion-extension was an average of 106.4 deg., for abduction-adduction was 73.6 deg., and for internal/external rotation was 106.6 deg. These values were reported to be higher than the ISO minimum values of 80 deg. for flexion-extension, 60 deg. for abduction-adduction, and 90 deg. for internal/external rotation per ISO 21535-2002.

Stress Analysis

While the BHR has been designed to minimize necessary bone resection of both the acetabulum and femoral head, the component design and materials utilized maintain sufficient material volume to withstand potential forces. Both instrumentation and surgical techniques are developed to allow checking of all cuts, etc., and to make appropriate corrections as necessary. The potential for surgical error, resulting in increase implant stress, is minimized by the policies on product labeling, operative technique and training of user surgeons on use of the device.

Static strength testing showed the average maximum load for dysplasia/bridging fixation screws was 88.4lbf (393 N) and 1099.6 lbf (4890 N) for the dysplasia/bridging cup flange; concluding the loads measured during testing under worst-case conditions indicate both components should be able to withstand predicted *in vivo* loads. The femoral head demonstrated an average yield point of 5620 N (1263 lbf). Fatigue testing of five femoral stems at a load of 143lbf for 5 million cycles showed no deformation or cracking. The results of these static tests demonstrate that the BHR components should withstand predicted *in vivo* loads.

Evaluation of equatorial roundness after simulated implantation

The equatorial roundness of the cup was evaluated after finishing, after insertion of cables and impaction into a balsa wood model. The equatorial roundness was 4.9micrometers after finishing, 3.4micrometers after insertion of cables, and 4.3micrometers after impaction. The impaction of the cup into balsa wood model appeared to have no detrimental effect on the equatorial roundness.

Acetabular Shell Beaded Surface and HA Coating Characterization

All acetabular cups have a single layer of integrally-cast CoCr-alloy (ASTM F75 and ISO 5832-4) beads on the outer surface that are coated with hydroxyapatite (HA) (ASTM F1185).

Microstructural analyses and bench testing were performed to characterize the beaded surface and the HA coating. The average total thickness of the beaded surface with HA coating was 931 micrometers. The average pore size (mean void intercept length) was 712 micrometers and the average % porosity (volume percent of void) was 34% for the beaded surface with HA coating.

Chemical and crystallographic analyses including XRD and IR spectra were provided to characterize the HA powder and coating materials. In addition, environmental stability of the HA coating on the fixation surface of the BHR acetabular cup was provided which characterized the

average solubility product (Ksp) of the plasma sprayed HA coating and the dissolution rate of the plasma sprayed HA powder.

The static shear strength of the beaded surface was 46MPa (6658psi). The static tensile strength of the beaded surface was 43MPa (6182psi). The results of the static shear and tensile strength tests surpassed the recommended value of 20 MPa (2900 psi) in the FDA *Guidance Document for Testing Orthopedic Implants with Modified Metallic Surfaces Apposing Bone or Bone Cement.* The shear fatigue strength of the beaded surface was 11MPa (1595psi). For all of these tests, ASTM methods were used, all failures were due to debonding epoxy, and no beads failed during testing. Because the beaded surface is cast with the substrate and not sintered, it is integral with the substrate. Therefore, the abrasion resistance should be equivalent to currently available CoCr porous beaded coatings.

The average shear strength for the HA coating was 30.5MPa. The average tensile strength for the HA coating was 12.8MPa. All failures were a result of debonding HA from the surface of the test specimen.

This characterization demonstrates that the beaded surface and HA coating has adequate strength and physical properties to perform as it is intended.

Explant Analysis

Two explant analysis reports for the BHR titled "Wear Retrieval Analysis of Birmingham Resurfacing," and "Finsbury Test Report "FI98001" were provided.

In the first report, the wear characteristics of the BHR device were investigated using 3 pairs of BHR bearings that were explanted from patients at 6-18mo post-implant. The 3 BHR devices were retrieved for the following reasons: (1) patient died of unrelated causes; (2) avascular necrosis of the femoral head; and (3) infection. These patients were known to be active for at least 6 months after receiving the device. An instrument with a resolution of 0.01micrometers was used and no measurable wear was detected as compared to their manufactured form.

In the "Finsbury Test Report, FI98001" an explant analysis of one BHR device was performed to evaluate roundness and surface finish. The device was retrieved following femoral neck fracture. The device was in place for 4 months. Results showed that the BHR head diameter was unchanged. Roundness was changed slightly equatorial from 0.3μm to 4μm and polar from 3.9μm to 6.5μm. Cup roundness increased on the equatorial from 0.6μm to 3.4μm. The report stated that the head showed approximately 80% as slightly dulled or "worn;" but, the authors reported that there was no undue damage or abnormalities. Surface finish changes were not significant.

The analysis of these 4 retrieved BHR devices demonstrated that they had undergone little change up to 18-months post-implantation.

Sterilization

Sterilization of the BHR components is in conformance with the following standards:

- BS EN 552:1994, *Sterilization of Medical Devices - Validation and Routine Control of Sterilization by Irradiation*;
- EN 46002:1997, *Quality Systems - Medical Devices - Particular Requirements for the Application of BS EN ISO 9002*;

- ISO 11137:1995, *Sterilization of health care products - Requirements for validation & routine control - Radiation & Sterilization*

The principal sterilization process is the Gamma Irradiation utilizing Cobalt 60, dose of 25-35kGy. The sterility assurance level (SAL) is 10^{-6}. The product is not labeled "pyrogen free".

Shelf Life Testing

The BHR components are packaged in a Tyvek™ vacuum peel pouch to maintain sterility. Shelf life testing was performed to verify sterile packaging integrity equivalent to five years.

X. SUMMARY OF CLINICAL STUDIES

INTRODUCTION:

A clinical data series was used to support the safety and effectiveness of the Birmingham Hip Resurfacing (BHR) system. The BHR was implanted in 2,385 hips by a single investigator, Mr. Derek J.W. McMinn, FRCS. Mr McMinn performed his surgeries at the Birmingham Nuffield and Little Aston Hospitals, Birmingham, United Kingdom from July 1997 through May 2004. Additionally, unpublished data on 3,374 hips implanted by 140 surgeons and published reports from the experience of multiple surgeons implanting over 3,800 hips supported the safety and effectiveness of the BHR System.

STUDY OBJECTIVES AND ASSESSMENTS:

The objective of the clinical data series was to demonstrate the safety and effectiveness of the Birmingham Hip Resurfacing (BHR) System. The safety assessments included data on revisions, adverse events, deaths and a metal ion literature review. The effectiveness assessments included survivorship and radiographic data, pain and function data as evaluated by the Oswestry-modified Harris Hip (OSHIP) Score, and patient satisfaction data.

PATIENT SELECTION METHODS AND INDICATIONS FOR USE:

In the case series, patients were not enrolled by the single investigator for pre-defined conditions; instead, a list of diagnoses for the BHR patients was provided. It was noted that during the same time period the investigator implanted the BHR devices, he also had patients who either had no surgery or a total hip replacement (THR). Therefore, to retrospectively develop the indications for use and physician labeling from the experience gained from the 2,385 implantations, a list of the factors that contributed to the investigator's decision to perform a THR in certain patients rather than the BHR hip resurfacing procedure was provided. These factors included:

- Advanced age: Patients of advanced age, especially those with low activity levels, were typically candidates for THR rather than BHR. Only 8.1% of the 2,385 cases included in the Overall McMinn cohort were >65 years of age. In these cases, BHR was selected despite advanced age if the patients had high activity levels, and had good bone stock of the femoral head.

- Low activity level: Patients with a low activity level were considered at lowered risk for future revision, and therefore good candidates for THR. Low activity level was characterized by no participation in sports activities, no heavy work required by job, a sedentary/retired lifestyle, or comorbidities that precluded a high activity level, such as severe arthritis in other joints or severe heart disease.

- Poor bone stock: Patient with poor bone stock were selected for THR rather than BHR because they were considered at risk for femoral neck fracture or femoral head collapse with a hip resurfacing procedure. Poor bone stock was characterized as severe osteopenia of the femoral head or femoral neck (determined by risk factors, medical history and/or diagnostic imaging), extensive AVN (>50% of femoral head, regardless of FICAT Grade), or the presence of multiple cysts.

The investigator's collection of a patient's pre-operative history, physical, and diagnostic work-up was commonly sufficient to screen candidates for BHR versus THR, and that only in rare instances would the planned surgical procedure be revised intraoperative. Although the investigator rarely changed his preoperative plan based on intraoperative findings, all patients were consented for a hip arthroplasty, and informed about the probable type of prosthesis they would receive. As with any surgical procedure, patients were also informed that based on the intraoperative findings, there could be changes to the planned procedure. The patients were thus consented for both a BHR and THR procedure. Based upon the 2,385 procedures studied, the factors outline above, and an analysis of the BHR revisions in the Overall McMinn Cohort (i.e., femoral neck fracture, femoral head collapse, dislocation, AVN, and infection), the indications for use and physician labeling were developed.

In addition to the factors described above, the literature findings were also considered. For example, the review of 50 BHR femoral neck fractures reported by Shimmin and Back[1] aided in the development of the labeling. In this publication, the authors reported on a review of 3,497 BHR cases performed in Australia by 89 surgeons. There were 50 femoral neck fractures in the series (or 1.46%) which the authors attributed to osteoporosis, and difficulties in implantation of the head and cup leading to notching of the superior femoral neck, varus placement of the device by more than 5°, difficulty in intra-operative alignment, impaction of the femoral component, and poor exposure. Based on these findings, warnings and precautions in the labeling that address these device related failures were included.

DESCRIPTION OF COHORTS AND DATA COLLECTED

The 2,385 procedures implanted with the Birmingham Hip Resurfacing (BHR) device by a single investigator from July 1997 through May 2004 were divided into the following three main cohorts for the purposes of data analysis:

- **X-ray cohort:** First 124 BHR cases performed from July 1997 through December 1997.
- **Oswestry cohort:** Next 1502 BHR cases performed from January 1998 through March 2002.
- **McMinn cohort:** Next 759 BHR cases performed from April 2002 through May 2004.

Table 3 outlines the dates of implantation, number of procedures, and types of safety and effectiveness data collected for these 3 cohorts:

Table 3: Cohorts and Data Collected			Types of Safety and Effectiveness Data Collected						
			Safety Data Collected			Effectiveness Data Collected			
Cohort	Dates of Implantation	Number of Procedures	Adverse Events	Revisions	Deaths	Survivorship	Radiographic	Pain and Function (OSHIP)	Patient Satisfaction
X-ray	7/97-12/97	124	X	X	X	X	**X**	X**	X
Oswestry	1/98-3/02	1502	X	X	X	X		X**	X
McMinn	4/02-5/04*	759*	X	X	X	X		***	

[1] Shimmin AJ, Back D. *Femoral neck fractures following Birmingham hip resurfacing: A national review of 50 cases.* J Bone Joint Surg [Br] 87-B:463-4, 2005.

Note: An X in the table indicates that this data was collected for the respective cohort

* There were 5 cases in the McMinn cohort whose implantations were performed prior to 4/02. These cases should have been part of the Oswestry cohort, but for unknown reasons were not. Therefore, unlike the majority of the McMinn cohort, some of these 5 cases have longer term follow-up.

** See note in Table 4 below regarding the number of procedures contributing to the pain and function (OSHIP) effectiveness data.

*** The pain and function data for the procedures in the McMinn cohort were collected using the Oxford Hip Score evaluation method (and not the OSHIP Score assessment method). Because the 759 procedures in the McMinn Cohort were not tracked by the Oswestry Outcome Center but by the National Health Services (NHS) Center, the FDA and Smith & Nephew, Inc. did not have access to the Oxford hip score data.

As noted in the Table above (with the large bolded "X"), 124 procedures in the X-ray cohort contributed to the assessment of radiographic effectiveness in the PMA. Radiographic evaluations were not provided for the 1502 procedures in the Oswestry cohort or the 759 procedures in the McMinn cohort.

Where there were common data elements collected in the 3 cohorts outlined above, this information was pooled into the following two combined cohorts:

- **X-ray/Oswestry/McMinn combined cohort or Overall McMinn cohort:** Note that for the rest of this document, this cohort will be referred to as the **Overall McMinn cohort.**
- **X-ray/Oswestry combined cohort**

Table 4 outlines the dates of implantation, number of procedures, and types of safety and effectiveness data collected for these 2 combined cohorts:

Table 4: Combined Cohorts and Data Collected			Types of Safety and Effectiveness Data Collected						
			Safety Data Collected			Effectiveness Data Collected			
Cohort	Dates of Implantation	Number of Procedures	Adverse Events	Revisions	Deaths	Survivorship	Radiographic	Pain and Function (OSHIP)	Patient Satisfaction
Overall McMinn Cohort	7/97-5/04	2,385	**X**	**X**	**X**	X	*	*	*
X-ray/ Oswestry Combined	7/97-3/02	1,626	X	X	X	**X**	*	**X****	**X**

Note: An X in the table indicates that this data was collected for the respective cohort

* Although data (e.g., x-ray or pain and function) was collected for one of the cohorts identified in this row, it was not collected for all procedures in the combined cohort; therefore, an X is not included in this part of the table.

** 1,111 unilateral procedures in the X-ray/Oswestry combined cohort contributed to the assessment of pain and function effectiveness data, as evaluated by the Oswestry-modified Harris Hip (OSHIP) Score assessment method.

As noted in the Table above (with large bolded "X"s), the 2,385 procedures in the Overall McMinn cohort contributed to the assessment of safety including adverse events, revisions, and deaths. The 1,626 procedures in the X-ray/Oswestry combined cohort contributed to the assessment of survivorship. Also, as noted in the Table above, 1,111 unilateral procedures in the X-ray/Oswestry combined cohort contributed to the assessment of pain and function effectiveness data, as evaluated by the Oswestry-modified Harris Hip (OSHIP) Score. Unilateral procedures

were evaluated separately as it is difficult to distinguish pain and function status of each hip separately in patients with bilateral hip involvement. Finally, 1,626 procedures in the X-ray/Oswestry Combined cohort contributed to the patient satisfaction effectiveness.

Additional Data Sources:

In addition to the clinical data series cohorts, less complete data was provided on 3,374 BHR cases performed by 140 surgeons worldwide (other than the single investigator). The follow-up for these cases was also contracted to the Oswestry Outcomes Centre and includes primarily the same parameters as the follow up for the X-ray/Oswestry combined cohort (adverse events, revisions, deaths, pain and function (OSHIP) scores, and patient satisfaction). The Oswestry Outcomes Centre, therefore, collected data on a total of 5,000 BHR cases. These 5,000 cases are referred to as the **Oswestry Worldwide Cohort**. The Oswestry Worldwide Cohort consists of 1) the 1,626 cases of the X-ray/Oswestry cohort (the single investigator), and 2) an additional 3,374 non-McMinn ("all other") cases. The Oswestry Outcomes Centre has provided access to all available data for the BHR cases from its database. Although the data from the 3,374 "all other" cohort was of some value, the applicant and FDA have no ability to independently verify any of the data provided to the Oswestry Outcomes Centre by sites other than the McMinn Center, and have no ability to request additional follow-up or clarifications of any kind from non-McMinn patients or physicians. For these reasons, the analysis on the Oswestry Outcomes Centre worldwide database has some limitations, and is not considered the primary data source.

Several literature references were also included which described the use of over 3,800 BHR devices implanted by multiple surgeons in several countries around the world. One example is the literature reference by Shimmin and Back, discussed above, which was used in the development of the labeling.

DATA COLLECTION METHODS

Safety Data Collection Methods

The safety data including adverse events, revisions, and deaths were collected by:

- The Oswestry Outcomes Center using an annual, patient-completed, mail-in questionnaire (deaths were identified while attempting to perform scheduled follow-up);
- The McMinn Center by recording the findings of post-operative patient visits to the McMinn Center in patient records; and
- Recording information provided to the single investigator by primary care physicians.

A summary was provided of the investigator's patient follow-up procedures which included regular evaluations in the preoperative and postoperative time periods according to standard practice and the investigator performed all revision surgeries (except in one known case). Therefore, the revision status was directly known to the investigator.

In addition, a summary was provided of the Oswestry Outcome Center's (OOC) follow-up procedures which included a collection of data on revisions and adverse events using an annual, patient-completed, mail-in questionnaire. Of the 180 cases missing their last theoretical expected mail-in questionnaire follow-up, 84 are missing at least 2 yearly evaluations, while 96 are only missing their last evaluation. These cases represent only 11.1% (180/1626) of the cases in the Oswestry/X-Ray Combined Cohort. With the exception of 8 cases classified by OOC as "no consent" (subjects who withdrew or did not agree to participate in the study), all other cases are not considered lost-to-follow-up by OOC since they continue to make attempts to contact patients.

Also, a 100% audit of all 2,385 procedures in the Overall McMinn Cohort was performed.

In addition to the safety data collection methods outlined above, a metal ion literature analysis was provided. Included in the analysis was an unpublished report by Daniel J, Ziaee H, and McMinn D, entitled, "Metal ion studies in patients treated with the Birmingham Hip Resurfacing, a comparable marketed device and historic metal-metal total hip replacements." The authors conducted 4 metal ion studies in patients who received BHR, Metasul metal-metal total hip replacements, and other marketed (historic) metal-metal total hip replacements.

In addition, summary of several literature references pertaining to the medium and long-term safety of cobalt and chromium ion exposure in the subject device (BHR), metal-on-metal total hip replacements, and metal-on-polyethylene total hip replacements was provided. These literature references are summarized below in the Summary of Safety Data.

Effectiveness Data Collection Methods

Survivorship Data Collection Method:

The primary effectiveness measurement was the X-Ray/Oswestry combined cohort survivorship study that included 1626 procedures performed from July 1997 through March 2002 at the Birmingham Nuffield Hospital. These procedures were a minimum of 2 years post-op. Of the 1626 procedures, data are available for 546 of the 601 BHR procedures eligible for 5-year follow up (90.8%). The data for the survivorship study was collected using the same methods presented above for the safety data collection methods.

Radiographic Data Collection Method:

The clinical data used to support this series contained the results of an independent radiographic review of the X-Ray Cohort, the first 124 procedures performed in the series from July 1997 through December 1997. Radiographic evaluations were not provided for the 1502 procedures in the Oswestry Cohort or the 759 procedures in the McMinn Cohort.

Baseline films for the purposes of comparisons were made in each of the 108 cases in the postoperative time period (usually within 3 months, but 8 of the 108 procedures had baseline evaluations performed at time points ranging from 110-860 days).

The radiographs were interpreted by an independent radiologist. A prospective protocol was developed and used to assess the radiographs. The 5-year anterior/posterior (AP) and lateral view radiographs were compared with the baseline radiographs for the medial-lateral migration, acetabular orientation (tilt angle), femoral and acetabular radiolucencies, heterotopic ossification (HO), bone resorption, acetabular protrusion, cysts, buttressing, and other abnormalities. Radiolucency was defined as a lucent area parallel to and in close proximity to the prosthesis/bone interface encompassing at least 50% of the zone and at least 1mm in width.

A radiographic success was defined as having all of the following:

- Absence of radiolucencies or a radiolucency in any one or two zones (a score of 0-6);
- Component migration $\leq$2mm; and
- Change in acetabular angle $<5°$

A radiographic failure was defined as the following:

- Presence of incomplete or complete radiolucencies or a radiolucency in all zones (a score of 7 or 8);
- A migration of the component >2mm; or
- A change in acetabular orientation of ≥5°

The individual success criterion was the absence of radiographic findings that suggest revision is necessary.

Oswestry-Modified Harris Hip (OSHIP) Score Data Collection Method

The clinical data used to support this series were collected by the Oswestry Outcomes Center (OOC) using an annual, patient-completed, mail-in questionnaire. The responses to the pain, function, and movement questions in the questionnaire were used to generate the Oswestry-modified Harris Hip (OSHIP) Score.

The OSHIP questionnaire allows patient assessments without direct physician or examiner evaluation. To address how the OSHIP data was collected, the OOC's standard operating procedures for data input and clarification were summarized for the patient-administered OSHIP questionnaires:

- Any questionnaires with missing, unclear. or conflicting information were returned to the patients with specific instructions on completing the form. The preferred method of follow-up was by mail; however, e-mail and telephone were also used.
- If the data were not collected, the score for any missing item was assumed to be the lowest possible (typically zero).

To address how the OSHIP scoring system was developed, an unpublished paper titled, "A Self-completed Tool for Evaluation of Hip Function: The Oswestry Hip Score," D. Barnes and co-workers reported that the OSHIP was developed by Professor James Richardson FRCS (Orth), Professor of Orthopaedics at the Institute of Orthopaedics, Robert Jones and Agnes Hunt Orthopaedics and District Hospital—NHS Trust in Oswestry, Shropshire, England. According to Barnes' paper, creation of the OSHIP began with the following premises:

- Long-term evaluation following hip replacement is essential, follow-up must be regular; and large-samples are necessary.
- Long-term and large-sample follow-up is difficult to obtain when using a score that requires surgeon- or radiologist-assessment.
- Physician-administered surveys are susceptible to bias (which may inflate the final scores) and may not truly represent the patients' own feelings; and
- Questionnaires needed to be simple and relatively short to make long-term and large-scale collection of data more efficient.

Building on these premises, Professor Richardson developed the OSHIP by combining elements of both the Harris and Merle d'Aubigne scores. The OSHIP produces an overall index score similar to that of the Harris score between 0 (worst) and 100 (best). Both the OSHIP and Harris Hip Score (HHS) are made up of the three domains of pain, function, and hip movement, with function being further divided into gait (walking, limp, and distance), and activity (stairs, sitting and transport).

The main difference between the OSHIP questionnaire and the HHS is that the OSHIP allows patient assessments without direct physician or examiner evaluation. In addition, the OSHIP questionnaire does not include the three HHS questions regarding physician assessment of Range

of Motion (5 pts.), Absence of Deformity (4 pts.), and the patient's ability to put on socks/tie shoes (4 pts.) but substitutes a "movement" question (13 pts.) that is intended for the patient to estimate their ability to flex their hip.

To justify the use of the OSHIP scoring system and the validity of patient self-administered questionnaires, several literature references were summarized.

Several researchers have reported a close correlation between patient self-assessment and physician assessment. Research by Mahomed et al.[2] demonstrated that patients are able to accurately respond to Harris Hip Score questions regarding pain and function with little difficulty, and that there is excellent correlation between the overall HHS pain/function scores reported by patients and the overall HHS pain/function scores reported by physicians (with a correlation of $r=0.99$, $p<0.0001$). Note that the Mahomed study did not include patient or physician evaluations of range of motion or deformity, these questions were eliminated from both the patient and physician assessments. Furthermore, McGrory et al.[3] found that a brief follow-up phone call (similar to the OOC follow-up procedure discussed above) was effective in capturing missing data and clarifying multiple or contradictory responses from mailed patient self-assessment questionnaires.

In addition, Barnes et al. evaluated the reliability and validity of the Oswestry Hip Score as documented in the research paper, "A Self-completed Tool for Evaluation of Hip Function: The Oswestry Hip Score." When evaluating the reproducibility of responses to two OSHIP questionnaires completed two weeks apart by 61 patients, the total intra-class correlation coefficient was 0.93 with intra-class correlation coefficients for the individual items and domains ranging from 0.67 to 0.92.

The correlation between the patients' overall self-administered OSHIP scores and physiotherapist-administered overall HHS scores in 28 consecutive patients was 0.91 ($p<0.0001$). Correlation between the individual corresponding domains of the Oswestry Hip Score and Harris Hip Score ranged from 0.60 and 0.89. The strongest correlation was between the domains of 'stairs' and 'walking/support' (0.89) and the lowest for the domains of 'limp' (0.60). Additional correlations not included in Mr. Barnes' study were provided. Correlation between the OSHIP "movement" domain and the HHS "shoes & socks," "deformity," and "range of motion" domains were performed. The correlation between OSHIP "movement" and HHS "shoes and socks," and HHS "range of motion," was 0.40 and 0.21, respectively. The correlation between OSHIP "movement" and HHS "deformity" was not included and not useful because all 28 subjects scored the maximum of 4 points on the HHS scale (score is either 0 or 4). Additional correlations were performed between OSHIP "movement" domain and the sum of the scores for the HHS "range of motion," "shoes and socks," and "deformity." The correlation between these items was calculated because the OSHIP "movement" domain is the substitute for the HHS "range of motion," "shoes and socks," and "deformity" domains. The correlation was calculated to be 0.40. In addition, a linear regression analysis was performed to predict HHS total score from OSHIP total score for the 28 subjects. The linear regression analysis calculated R^2 is approximately 0.83, which measures the proportion of total variation about the mean explained by the linear regression model. Due to an unclear randomization scheme and questionable masking procedure used to select these 28 sample patients, it is not

---

[2] Mahomed NN, Arndt DC, McGrory BJ, Harris WH. *The Harris Hip Score: Comparison of patient self-report with surgeon assessment.* J Arthroplasty 16(5):575-80, 2001.

[3] McGrory BJ, Shinar AA, Freiberg AA, Harris WH. *Enhancement of the value of hip questionnaires by telephone follow-up evaluation.* J Arthroplasty 12(3), 1997.

easy to generalize the above correlations to the general target patient population and clinical judgment was sought from the Orthopaedic and Rehabilitation Devices Advisory Panel.

A review of the raw data from the 28 patient Barnes' study revealed the following:

- The average OSHIP score was lower than the HHS score, 62 and 67, respectively.
- Fewer subjects had an OSHIP score greater than 80 and more subjects had an OSHIP score less than 70 as compared to their HHS score.
- There were 14 pairs of data where the OSHIP and HHS scores differed by more than 5 points. Of the 14 pairs, the HHS score was higher in 12 cases while the OSHIP was higher in only 2 cases.

Therefore, Barnes emphasized the tendency of the OSHIP scores to be somewhat lower relative to the HHS scores, suggests that the OSHIP is a very close, although *conservative*, estimate of the HHS.

A paper by Ragab[4] and co-workers reported a lack of correlation between patient self-assessment of pain and function and physician assessment of pain and function (with a correlation of $r=0.467$, $p<0.01$), Like the Barnes study, Ragab[5] also reported a relative lack of correlation between patient assessment of limp and physician assessment of limp which he believed was due to the physician's tendency not to report limps that occurred only after long walks or during weather change, while patients were likely to report such limps. However, unlike the Barnes study in which the OSHIP and HHS item regarding "pain" had a correlation of 0.83, Ragab found that when the patients reported significant pain, they were often attributing the pain to their hips when the pain, in most cases, was not truly hip related. The author reported that the physician was better able to distinguish "true" hip pain from pain coming from other sources (for examples, secondary to trochanteric bursitis, lumbar spondylosis, and arthrosis of the contralateral hip). Although Ragab concluded that there is a lack of correlation between patient and physician assessments, Ragab's research does confirm Barnes' findings that patient self-assessments tend to be lower than physician assessments.

An additional finding by McGrory and co-workers[6] was that questions about whether patients could cut their toenails and put on socks/shoes correlated significantly with the HHS range of motion calculation with correlations of $r=0.57$ and $r=0.53$, respectively. The authors concluded that responses to these two questions could therefore be used to estimate the weighted HHS range of motion. In addition, Johnston and Smidt[7] reported that there is a distinct relationship between hip flexion and shoe tying. These articles suggest that an evaluation of a patient's movements during specific activities of daily living correlate well and may substitute for a physician evaluation of ROM as outlined in the HHS.

---

[4] Ragab A.A. *Validity of self-assessment outcome questionnaires: patient-physician discrepancy in outcome interpretation*, Biomed Sci Instrum (2003); 39: pp.579-84

[5] Ragab A.A. *Validity of self-assessment outcome questionnaires: patient-physician discrepancy in outcome interpretation*, Biomed Sci Instrum (2003); 39: pp.579-84.

[6] McGrory BJ, Freiberg A, Shinar AA, Harris WH. *Correlation of measured range of motion following total hip arthroplasty and responses to a questionnaire*. J Arthroplasty 11(5):565-71, 1996.

[7] Johnson RC, Smidt GL. *Hip motion measurements for selected activities of daily living*. Clin Orthop 72:205-216, 1970.

Patient Satisfaction Data Collection Method

Patient satisfaction data were also collected using the annual, patient-completed, mail-in questionnaire. For the purpose of the BHR study, an additional question about patient satisfaction was appended to the end of the OSHIP assessment questionnaire.

LITERATURE REFERENCES:

A literature search was performed to find published studies of ceramic-on-ceramic total hip replacements to provide a comparison for the BHR clinical study data. PaperChase internet service was used to conduct the literature search and found 400 citations. The abstracts were reviewed and excluded if the article was not in English; was conducted prior to 1990; was a review article; was a small case series with <25 patients; had a highly select patient population; had no specific device identification available; did not use the Harris Hip Score; and did not have a 2-year minimum follow-up. Only two literature articles met these criteria:

D'Antonio J., et al.: New experience with alumina-on-alumina ceramic bearings for total hip arthroplasty. J. Arthroplasty, 17(4): 2002.

This clinical dataset is the same group of 514 procedures that are included in the Howmedica Osteonics ABC System and Trident System PMA (P000013) that used a CoCr alloy femoral stem and a porous-coated Ti alloy acetabular shell with Alumina Bearing Couple (ABC) and the hydroxyapatite-coated titanium shell.

Garino JP: Modern ceramic-on-ceramic total hip systems in the United States: Early results. Clin. Orthop., 379: 2000.

This clinical dataset is the same group of 333 procedures presented in Wright Medical's Ceramic Transcend Articulation System PMA (P030027).

The data in these references have some differences as compared to the data provided for the BHR device in this clinical data series, including

- Different evaluations, (OSHIP for BHR and HHS for literature)
- Length of follow-up, (18-36mo and 2-4 years for the controls and 2-5 years for the BHR study)
- Mean baseline pain and function scores (e.g., 60 for OSHIP in BHR Oswestry cohort, 44 for HHS Garino study, and not reported for D'Antonio study), and
- Indications for use, (including differences in the rate of dysplasia and AVN diagnostic indications)

However, the literature information provided valuable information on approved ceramic-on-ceramic THR systems for comparison purposes including patient demographics, diagnostic indications, patient accounting, adverse events, revision rates, pain, function, and radiographic results. This information is summarized in several sections below for reference purposes.

PATIENT DEMOGRAPHICS

Demographics for X-Ray, Oswestry, McMinn, and Overall McMinn cohorts

Patients in the Overall McMinn cohort were 70.6% men and 29.4% women, ages 13-86 years (average 53.1 years). The primary diagnosis was osteoarthritis in 75.0%, dysplasia in 15.8%, avascular necrosis in 4.1%, inflammatory arthritis in 2.4%, and "other" in 2.7% (Table 5).

Table 5: Procedure Demographics				
	X-Ray Cohort	Oswestry Cohort	McMinn Cohort	Overall McMinn
Hips	124	1502	759	2385
Men	81 (65.3%)	1082 (72.0%)	520 (68.5%)	1683 (70.6%)
Women	43 (34.7%)	420 (28.0%)	239 (31.5%)	702 (29.4%)
Age (range)	52.8 (27.8-75.3)	53.0 (13.4-86.5)	53.3 (21.6-79.5)	53.1 (13.4-86.5)
Age ≤65 years	111 (89.5%)	1388 (92.4%)	692 (91.2%)	2191 (91.9%)
Dx: OA	92 (74.2%)	1171 (78.0%)	526 (69.3%)	1789 (75.0%)
Dx: DDH	22 (17.7%)	197 (13.1%)	158 (20.8%)	377 (15.8%)
Dx: AVN	7 (5.6%)	59 (3.9%)	31 (4.1%)	97 (4.1%)
Dx: Inflammatory	2 (1.6%)	39 (2.6%)	16 (2.1%)	57 (2.4%)
Dx. Other	1 (0.8%)	36 (2.4%)	28 (3.7%)	65 (2.7%)

Demographics for X-Ray/Oswestry combined cohort

Patients in the survivorship study (X-ray/Oswestry combined cohort) ranged in age from 13.4 to 86.5 years (mean 53 years); 72% of the patients are male, and 28% are female. Of the 1,626 BHR procedures in this cohort, 1,499 (92%) were performed in patients ≤ 65 years old, and 127 (8%) were performed in patients > 65 years old.

Diagnostic Indications for Unilateral and Bilateral procedures in X-Ray/Oswestry combined cohort

One thousand one hundred and eleven (1,111) of the X-ray/Oswestry combined cohort cases (68%) were unilateral procedures and 515 (32%) were bilateral procedures. The indication for the majority of cases was osteoarthritis. Table 6 provides the breakdown of unilateral and bilateral cases by indication.

Table 6: Diagnostic Indication for BHR			
Diagnosis	Unilateral	Bilateral	TOTAL
Osteoarthritis	849 (76.4%)	414 (80.4%)	1263 (77.7%)
Dysplasia	160 (14.4%)	59 (11.5%)	219 (13.5%)
Avascular necrosis	52 (4.7%)	14 (2.7%)	66 (4.1%)
Inflammatory arthritis	18 (1.6%)	23 (4.5%)	41 (2.4%)
Other	32 (2.9%)	5 (1.0%)	37 (2.3%)
TOTAL	1111 (68%)	515 (32%)	1626

Some of the patients with bilateral hip replacements were included in different groups depending on when the second hip procedure was performed (Table 7).

Table 7: Hip Procedures								
Cohort	Patients**	Hips***	Unilateral	Bilateral	Contralateral Single Hip Cohort*			Singles
					X-Ray	Oswestry	McMinn	
X-Ray	113	124	83	11	-	11	8	19
Oswestry	1301	1502	1028	201	11	-	61	72
McMinn	685	759	542	74	8	61	-	69

* Patients with bilateral hip replacements with the contralateral hip not included in the first hip replacement's evaluation cohort.

** Number of patients equals unilateral + bilateral + singles
*** Number of hips equals unilateral + (2 x bilateral) + singles

Demographics: Literature References

The study published by D'Antonio *et al.* reported findings from a multicenter study conducted at 22 investigational sites; the study published by Garino was conducted at 11 investigational sites (Table 8).

Table 8: Demographics for Literature References

Author	Patients	Procedures	Age (Average)	Bilateral Procedures
D'Antonio J *et al*	458	514: • 349 ceramic • 165 control	53	19
Garino JP	333 (f=132, m=201)	333	52	0

D'Antonio *et al.* reported the indication for THR as osteoarthritis in 399/514 procedures (77.6%) and avascular necrosis in 82/514 procedures (16%) (Table 9).

Indication for Arthroplasty

Table 9: Indications for Use for Literature References

Diagnosis	D'Antonio
OSTEOARTHRITIS	399
TRAUMATIC OSTEOARTHRITIS / DJD	21
AVASCULAR NECROSIS	82
OTHER / NOT REPORTED	12
TOTAL	**514**

DESCRIPTION OF DEVICE IMPLANTATIONS

The following information on the femoral head sizes and acetabular cup styles and sizes implanted in the 2385 procedures in the Overall McMinn cohort was provided (Table 10).

Table 10: Surgical Data: Implant Sizes
All Patients

Acetabular Cup	Femoral Resurfacing Component/Head					
	38mm	42mm	46mm	50mm	54mm	58mm
44mm	2 (0.1%)					
46mm	5 (0.2%) 4[D] (0.2)					
48mm		119 (5.0%)				
50mm		67 (2.8%) 39[D] (1.6)				
52mm			342 (14.3%)			
54mm			154 (6.5%) 1[C] (0.0) 50[D] (2.1)			
56mm				683 (28.6%)		
58mm			3[B] (0.1%)	167 (7.0%) 28[D] (1.2)		

Table 10: Surgical Data: Implant Sizes All Patients						
Acetabular Cup	Femoral Resurfacing Component/Head					
	38mm	42mm	46mm	50mm	54mm	58mm
60mm					460 (19.3%)	
62mm				1 [B] (0.0)	137 (5.7%) 38 [D] (1.6)	
64mm						51 (2.1%)
66mm					1 [B] (0.0)	22 (0.9%) 10 [D] (0.4)

B Bridging cups
C Custom cups
D Dysplastic cups

Stratification of Results by Hybrid/Cement/Uncemented:
There was only one case (of the 1,626 cases in the X-ray/Oswestry combined cohort) in which the femoral component was not cemented (a customized implant to accommodate broken metal that remained in the femoral head from a previous event). Therefore, the number of non-hybrid implants (cemented femoral resurfacing component/uncemented acetabular cup) was negligible.

PATIENT ACCOUNTING

The follow-up rates for the Combined X-Ray / Oswestry Cohort, upon which the effectiveness analyses were performed, at the 1-year, 2-year, 3-year, 4-year, and 5-year postoperative evaluation time points were 76.6%, 77.3%, 88.1%, 88.6%, and 90.8%, respectively. There were 546 procedures (hips) evaluated at 5 years in this cohort (Table 11).

Table 11: Patient Accounting Based on the number procedures						
	Baseline	1 year	2 years	3 years	4 years	5 years
Accounting for Survivorship (% Revision Free)						
Cohort		# Patients observed at beginning of each study year (# revisions, # censored) [1]				
X-Ray	-	124 (1,0)	123 (0,0)	123 (1,0)	122 (0,0)	122 (0,20)[6]
Oswestry	-	1502 (9,63)	1430 (5,49)	1376 (4,256)	1116 (1,321)	794 (1,392)
McMinn	-	759 (3,290)	466 (0,379)	87 (0,84)	3 (0,0)[7]	3 (0,0)[7]
X-Ray Cohort						
Expected[1,8]	124	123	123	122	122	118[5]
Evaluated[2]	82	101	51	122	119	112
F/U %[2]	66.1%	82.1%	41.4%	100.0%	97.5%	94.9%[3]
Evaluated[4]	124	-	-	-	-	108
F/U%[4]	100%	-	-	-	-	91.5%
Oswestry Cohort						
Expected[1,8]	1502	1493	1484	1227	885	482
Evaluated[2]	1229	1137	1192	1067	773	434
F/U %[2]	81.8%	76.2%	80.3%	87.0%	87.3%	90.0%
X-ray / Oswestry Combined Cohort						
Theoretical[1]	1626	1626	1626	1385	1045	647
Deaths (procedures)	0	2	7	16	18	26
Revisions (cumulative)	0	10	15	20	21	23
Expected[1,8]	1626	1616	1607	1349	1007	601
Evaluated[2]	1311	1238	1243	1189	892	546
F/U %[2]	80.6%	76.6%	77.3%	88.1%	88.6%	90.8%

Table 11 Continued: Patient Accounting Based on the number procedures						
	Baseline	1 year	2 years	3 years	4 years	5 years
F/U +base[5]	1311	1067/1304	1050/1294	944/1046	660/726	368/397
+base %		82%	81%	90%	91%	93%
F/U –base[5]	315	171/312	193/313	245/303	232/281	178/204
-base %		55%	62%	81%	83%	87%

[1] Note that for the Survivorship data the "year 1" data is starting from day 1 and the "year 2" data is starting from day 366, etc. but for the OSHIP scores, the "year 1" data was collected between day 366-730, the "year 2" data was collected between day 731-1095, etc.
[2] Evaluated by OSHIP score
[3] OSHIP score was available for one hip that was revised shortly after the 5-year follow-up interval, OSHIP data available on 112/119 (94.1%) of hips surviving to 5 years
[4] Evaluated by X-Ray
[5] The follow-up of those who had baseline OSHIP scores (+base) and those without baseline OSHIP scores (-base).
[6] Note that there were 2 revisions in the x-ray cohort at >5 years
[7] There were 5 cases in the McMinn cohort whose implantations were performed prior to 4/02. These cases should have been part of the Oswestry cohort, but for unknown reasons were not. Therefore, unlike the majority of the McMinn cohort, some of these 5 cases have longer term follow-up.
[8] The expected and evaluated values in each interval include hips with a recorded OSHIP even if the subject died or was revised during the interval.

For the unilateral patients in the X-Ray / Oswestry combined cohort, the follow-up rates at the 1-year, 2-year, 3-year, 4-year, and 5-year postoperative evaluation time points were 75.7%, 76.6%, 88.2%, 88.4%, and 91.1%, respectively (Table 12).

Table 12: Patient Accounting Summary of the Oswestry and X-Ray Cohorts - Unilateral Based on Available OSHIP Data						
	Baseline	1 year	2 years	3 years	4 years	5+ years
Theoretical	1111	1103	1100	927	687	395
OSHIP data	892	835	842	818	607	360
%	80.3	75.7	76.5	88.2	88.4	91.1

Accounting identified in the literature references were as provided in Table 13.

Table 13: Patient Accounting: Literature References		
Author	**Mean follow-up (range)**	**Number of hips (patients) included**
D'Antonio	35.2 mos (24 to 48 mos) for ceramic on ceramic. 33.6 mo (24 to 48 mo) for control (metal on polyethylene)	349 ceramic-on-ceramic THR procedures (318 patients) • 335 hips (307 pts) at 24 mos • 243 hips (227 pts) at 36 mos • 72 hips (71 pts) at 48 mos 165 control THR procedures (161 patients), • 149 hips (147 pts) at 24 mos • 111 hips (111 pts) at 36 mos • 26 hips (26 pts) at 48 mos
Garino	Range 18-36 months	"100% follow up for all 333 procedures"

SAFETY DATA

Safety: Revisions

There were 27 procedures that required revision. Two of the 27 revisions occurred beyond the 5-year follow-up time point in the X-Ray cohort (Table 14).

Table 14: Revisions Stratified by Cohort

	X-Ray Cohort N=124					
	Preop	**1 year**	**2 years**	**3 years**	**4 years**	**5+ years**
Number of procedures*	124	124	123	123	122	122
Revisions	-	1	0	1	0	2
	Oswestry Cohort N=1502					
Number of procedures*	1502	1502	1430	1376	1116	794
Revisions	-	9	5	4	1	1
	McMinn Cohort N=759					
Number of procedures*	759	759	466	87	3	3
Revisions	-	3	0	0	0	0
	X-Ray + Oswestry Combined Cohort N=1626					
Number of procedures*	1626	1626	1553	1499	1238	916
Revisions	-	10	5	5	1	3
	Overall McMinn Cohort N=2385					
Number of procedures*	2385	2385	2019	1586	1241	919
Revisions	-	13	5	5	1	3

* The number of procedures is the number of hips that were surviving at the end of the previous year based on the survival analysis. Note that for the Survivorship data the "year 1" data is starting from day 1 and the "year 2" data is starting from day 366, etc.

There were 10 revisions due to a femoral neck fracture, 6 for femoral head collapse, 1 for dislocation, 2 for AVN (1 led to femoral head collapse and 1 led to a femoral neck fracture), and 8 for infections (2 led to head collapse, 1 led to a femoral neck fracture) (Table 15). Altogether, there were 12 femoral neck fractures that required revisions. Factors that may have contributed to the femoral neck fractures include age-related osteopenia (2 patients), poor preoperative bone quality as evidenced by cysts in the femoral head and acetabulum (1 case), systemic lupus erythematosus (SLE) (1 case), severe rheumatoid arthritis (RA) (1 case), infection that led to bone death (1 case), femoral head cysts (1 case), and malpositioned component (1 case). The 9 cases with femoral head collapse (6 primary femoral head collapses, 2 collapses due to infection and 1 due to AVN). Factors that may have contributed to the femoral head collapse include infection (2 cases), AVN (2 cases), femoral head cysts and soft bone (3 cases), osteopenia (1 case), and 1 unknown.

Table 15: Number and reason for revisions stratified by study cohort (all hips)						
Cohort	Infection	Femoral neck fracture	Collapse femoral head	Avascular Necrosis (AVN)	Dislocation	Mean in days (SD)
X-Ray	3/124* (2.4%)	1/124 (0.8%)	0/124	0/124	0/124	1252 days (848)
Oswestry	5/1502 (0.3%)	7/1502 (0.47%)	6/1502 (0.4%)	2/1502 (0.13%)	0/1502	495 days (466)
McMinn	0/759	2/759 (0.26%)	0/759	0/759	1/759 (0.13%)	58.3 days (72.6)
Total	8/2385 (0.3%)	10/2385 (0.4%)	6/2385 (0.25%)	2/2385 (0.08%)	1/2385 (0.04%)	
Mean in years	3.12 yrs	0.2 yrs	2.2 yrs	0.66 yrs	(1 day)	

* Two of the 3 revisions due to infections beyond 5-year follow-up

Safety: Revisions Comparison with Literature References

A comparison of the revision rates between the BHR study cohorts and the two literature reference groups was provided. The revision rate for the primary efficacy cohort was 1.47% at 5 years compared to 1.2%, 5.2%, and 1.2%, respectively, for the D'Antonio ceramic-ceramic, D'Antonio metal-poly, and Garino literature reference groups (Table 16).

Table 16: Revision Rate Comparisons								
	Cohort					Literature Reference Data		
	X-Ray	Oswestry	X-Ray/ Oswestry Combined	McMinn	Overall McMinn	D'Antonio C/C*	D'Antonio M/P*	Garino
N	124	1502	**1626**	759	2385	338	151	333
Revised	4	20	**24**	3	27	4	8	4
Rate %	3.2%	1.3%	**1.47%**	0.3%	1.13%	1.2%	5.2%	1.2%
f/u years	5	4	**4-5**	1	3	3	3	1-3

* Revision rates are based on a minimum of 2-year follow-up

Safety: Adverse Events

A time course distributions of adverse events was provided (Table 17). The Overall McMinn Cohort contains the X-Ray, Oswestry, and McMinn cohorts, and can be considered the safety cohort for this study. Also, presented below, is a table with the total number of adverse events in the Overall McMinn Cohort stratified by adverse event type and compared with Literature Reference Groups (Table 18).

Table 17: Adverse Events* Overall McMinn Cohort						
Adverse Event*	**Overall McMinn Cohort N=2385**					
	Postop	**1 year**	**2 years**	**3 years**	**4 years**	**5+ years**
Number of procedures	2385	2157	1667	1378	1018	620
Procedures with AE (%)	1126 (46.2%)	847 (39.3%)	155 (9.3%)	64 (4.6%)	34 (3.3%)	53 (8.5%)
AVN femoral head/neck	31 (1.3%)	2 (<0.1%)	1 (<0.1%)	0	0	1 (0.2%)
Femoral head collapse	7 (0.3%)	3 (0.1%)	3 (0.2%)	1 (<0.1%)	0	1 (0.2%)
Component migration/loosening	1 (<0.1%)	7 (0.3%)	8 (0.5%)	2 (0.1%)	0	1 (0.2%)
Femoral neck fracture	0	10 (0.5%)	0	2 (0.1%)	0	1 (0.2%)
Impingement	2 (<0.1%)	1 (<0.1%)	0	0	0	0
Infection (device related)	0	7 (0.3%)	3 (0.2%)	1 (<0.1%)	1 (<0.1%)	2 (0.3%)
Dislocation	0	5 (0.2%)	0	2 (0.1%)	0	2 (0.3%)
Cardiac event	15 (0.6%)	1 (<0.1%)	0	1 (<0.1%)	0	0
Hg drop	179 (7.5%)	2 (<0.1%)	0	0	0	0
Heterotopic Ossification	0	33 (1.5%)	19 (1.1%)	3 (0.2%)	1 (<0.1%)	3 (0.5%)
Hypotension	33 (1.4%)	4 (0.2%)	0	0	0	0
Limp	0	203 (9.4%)	4 (0.2%)	2 (0.1%)	0	1 (0.2%)
Event at implant site (clicking, etc.)	0	51 (2.4%)	14 (0.8%)	9 (0.7%)	1 (<0.1%)	3 (0.5%)
Reaction at incision site	8 (0.3%)	62 (2.9%)	1 (<0.1%)	1 (<0.1%)	0	2 (0.3%)
Other (see description below)	171 (7.2%)	121 (5.6%)	19 (1.1%)	7 (0.5%)	7 (0.7%)	5 (0.8%)
Thromboembolic event	3 (0.1%)	3 (0.1%)	0	0	0	0
Pain	26 (1.1%)	223 (10.3%)	76 (4.6%)	22 (1.6%)	20 (2.0%)	29 (4.7%)
Deep Vein Thrombosis	5 (0.2%)	1 (<0.1%)	2 (0.1%)	0	0	0
Infection (hip/procedure related)	28 (1.2%)	13 (0.6%)	0	0	0	0
Pneumonia	2 (<0.1%)	0	0	0	0	0
Fever	171 (7.2%)	1 (<0.1%)	1 (<0.1%)	0	0	0
X-ray report comment	0	23 (1.1%)	12 (0.7%)	7 (0.5%)	3 (0.3%)	7 (1.1%)
Stiffness, weakness, flexion deformity, restricted ROM	0	184 (8.5%)	11 (0.7%)	9 (0.7%)	3 (0.3%)	3 (0.5%)
Urinary	234 (9.8%)	1 (<0.1%)	0	0	0	0
Wound exudate	588 (24.7%)	1 (<0.1%)	0	0	0	0

* Time course of events shows the number and % of subjects with at least 1 complication of the specified type in the specified time period. Subjects may appear in more than one time period. Events without time information were not included in the table.

Safety: Adverse Events Overall McMinn Cohort and Comparison with Literature Reference Groups

The rate of wound exudates differs significantly between the two literature reference groups and the Overall McMinn Cohort, 3.4% and 1.4% versus 25%. It was reported that this is probably due to a difference in the definition or reporting requirements. There does not appear to be a correlation between wound exudates and superficial or deep infections. AVN of the femoral head

(1%), femoral head collapse (<1%), and femoral neck fracture (<1%), which are not possible in conventional total hip replacements, occurred at low rates. The "other" adverse events in the Overall McMinn Cohort included non-device and non-procedure related adverse events, such as dizzy spells, rashes, illnesses, ankle fracture, prostate cancer, or other pre-existing medical conditions.

Table 18: Comparison of Adverse Events
Overall McMinn Cohort vs. Literature Reference Groups

Adverse Event	Overall McMinn Cohort Totals*	Garino Reference	D'Antonio Reference N=349		
			ABC with porous	ABC with HA	Control M/P
Number of procedures	**2385**	**333**	**172**	**177**	**165**
Procedures with AE (%)	1669 (70%)				
Total AEs	2912				
AVN femoral head/neck	35 (1%)				
Femoral head collapse	15 (<1%)				
Component migration/loosening	21 (<1%)				
Femoral neck fracture	13 (<1%)				
Impingement	3 (<1%)				
Infection (device related)	15 (<1%)	1 (<1%)			
Dislocation	8 (<1%)	3 (1%)	2.3%	3.4%	4.2%
Radiological AE	-				
Femoral calcar fracture		3 (1%)			
Acet liner misplaced		2 (1%)			
Liner chipped insertion		3 (1%)	2.9%	2.3%	-
Acetabular migration		1 (<1%)			
Shell malposition		1 (<1%)			
Bursitis		1 (<1%)			
Cardiac event	21 (<1%)				
Femoral fracture			2.4%	1.2%	1.2%
Hg drop	182 (8%)				
Heterotopic Ossification	56 (2%)		2.9%	3.4%	6.1%
Hypotension	37 (2%)				
Limp	211 (9%)		2%	4%	3%
Event at implant site (clicking, etc.)	75 (3%)				
Reaction at incision site	74 (3%)				
Other (see above description)	328 (14%)				
Thromboembolic event	7 (<1%)	1 (<1%)			
Pain	367 (15%)	2 (<1%)	9%	8%	7%
Deep Vein Thrombosis	8 (<1%)				
Infection (hip/procedure related)	41 (2%)	1 (<1%)			
Pneumonia	2 (<1%)				
Fever	177 (7%)				
X-ray report comment	53 (2%)				
Stiffness	206 (9%)				
Urinary	235 (10%)				
Wound exudate	589 (25%)		3.4%	1.4%	

Table 18 Continued: Comparison of Adverse Events Overall McMinn Cohort vs. Literature Reference Groups					
Adverse Event	Overall McMinn Cohort Totals*	Garino Reference	D'Antonio Reference N=349		
			ABC with porous	ABC with HA	Control M/P
Not applicable (pre-existing condition)	3 (<1%)				
Foot-drop		1 (<1%)			
Vertebral fracture		1 (<1%)			
Other local complication		8 (2.5%)			

* Overall event count shows the number and % of subjects that had an event at any time. Subjects were counted only once for an event regardless of the number of times the event was recorded. Events without time information are included in this table.

Safety: Adverse Events - Discussion of Infections

The infections identified in the clinical data series were categorized, based on data collection procedures, as hip/procedure-related or device-related based on the time of occurrence. There were 41 infections associated with the index hip resurfacing procedure within 30 days of surgery and were thus categorized as hip/procedure-related. All of these events were wound exudates or wound infections that resolved with antibiotics. There were 15 infections that occurred more than 30 days after surgery and were thus categorized as device-related. Of these 15 infections, 6 required revisions and 9 "resolved with antibiotics." There were two patients who were revised for other indications (component migration and femoral neck fracture) who were found to be infected.

Infections that involve the prosthesis will not typically be successively treated with antibiotics alone. Therefore, it is unlikely the 41 "hip/procedure-related" infections and 9 "device-related" infections that were resolved with antibiotics were actually device-related infections. Therefore, these should be categorized in the hip/procedure-related category, probably as wound problems or superficial infections.

Safety: Adverse Events - Deaths

There were 20 patient deaths (26 procedures) in the Overall McMinn Cohort. It was stated that in no case was a death related to the BHR procedure. The causes were reported to be: 2 stroke, 4 cancer, 1 motor neuron disease, 1 esophageal cancer and pneumonia, 1 myocardial infarction, 1 suicide, 1 ruptured aorta, 1 carcinoma prostate with metastases, 1 unconfirmed – either diving accident or myocardial infarction, 7 unreported.

Safety: Metal Ion Literature Analysis

The literature references were provided to address concerns for metal ion release.

- An unpublished report by **Daniel J, Ziaee H, and McMinn D, entitled, "Metal ion studies in patients treated with the Birmingham Hip Resurfacing, a comparable FDA-approved device and historic metal-metal total hip replacements"** was provided.

  The authors conducted 4 metal ion studies in patients who received BHR, Metasul metal-metal total hip replacements, and other marketed (historic) metal-metal total hip replacements:

  1. A short-term longitudinal study of urinary Co and Cr levels in patients with the BHR and the Metasul metal-metal total hip replacement.

12-hour urine collections were obtained preoperatively and postoperatively at 5 days, 2 months, 6 months, 1 year and 2 years for 26 consecutive patients who underwent BHR. The inclusion criteria were unilateral end-stage arthritis; 50mm and 54mm femoral heads; no other implanted metallic devices; and no renal failure. A comparison group of 28 Metasul metal-metal total hip replacement patients operated on 1-3 years previously were studied. The metal ion analyses were performed using a High Resolution Inductively Coupled Plasma Mass Spectrometer (HRICPMS). The mean urinary Co output was 0.4μg/day, 4.0μg/day, 9.0μg/day, 19.2μg/day, 13.4μg/day, and 12.3μg/day for the preoperative, 5-day, 2-month, 6-month, 1-year and 2-year postoperative time points, respectively. A comparison was made between these values with the mean of 11.6μg/day in the 28 Metasul metal-metal total hip replacement patients at 1-3 years. The mean urinary Cr output was 1.6μg/day, 2.1μg/day, 4.0μg/day, 7.3μg/day, 5.3μg/day, and 5.3μg/day for the preoperative, 5-day, 2-month, 6-month, 1-year and 2-year postoperative time points, respectively. In addition, a comparison was made between these values with the mean of 3.7μg/day in the 28 Metasul metal-metal total hip replacement patients at 1-3 years.

2. A long-term cross-sectional study of urinary Co and Cr levels in patients with the BHR and the Metasul metal-metal total hip replacement.

12-hour urine collections were obtained from 58 patients who 5 years previously underwent BHR and 23 patients who received a Metasul metal-metal total hip replacement. At 5 years, the mean urinary Co output was 13.3μg/day for the BHR patients and 14.2μg/day for the Metasul patients. At the same time period, the mean urinary Cr output was 6.4μg/day for the BHR patients and 4.1μg/day for the Metasul patients

3. A longitudinal study of whole blood Co and Cr levels in patients with the BHR.

Whole blood samples were obtained preoperatively and 1 year postoperatively for 26 consecutive patients who underwent BHR (the same patients as the longitudinal study described above). In addition, the 58 patients who underwent BHR 5 years previously were also studied. The whole blood Co levels were 0.2μg/l, 1.3μg/l, and 1.8μg/l at the preoperative and 1-year time points and the 5-year study patients, respectively. The whole blood Cr levels were 0.3μg/l, 2.4μg/l, and 1.6μg/l at the preoperative and 1-year time points and the 5-year study patients, respectively.

4. A cross-sectional study of whole blood Co and Cr levels in patients with the BHR and metal-metal total hip replacements.

Whole blood samples were obtained 1 year postoperatively for 16 BHR patients who were described as "high quality sportspersons," i.e., very physically active. Whole blood samples from 20 patients who underwent Metasul metal-metal total hip replacements 1 year previously and 16 patients who had "historic" metal-metal total hip replacements (Ring and McKee Farrar) were also studied. The mean whole blood Co levels were 2.7μg/l in the sportspersons BHR group, 2.1μg/l in the historic metal-metal THR group, and mean whole blood Cr levels were 5.8 μg/l in the sportspersons BHR group and 3.4μg/l in the historic metal-metal THR group.

The authors compared the measured 1-year and 5-year BHR and 1-year Metasul whole blood Cr levels (2.4μg/l, 1.7μg/l and 1.6μg/l) with the 17μg/l "safe limit" as proposed by the EKA (Expositionaquivalente fur Krebserzeugende Arbeitsstoffe). The authors also compared the measured 6-month, 1-year and 2-year mean urinary output of Cr for the BHR patients

(4.07µg/g creatinine, 4.24 µg/g creatinine, and 4.89µg/g creatinine) with the 300µg/g creatinine Biological Exposure Index for Cr as recommended by the ACGIH (American Conference of Government Industrial Hygienists).

The daily urinary output of metal ions in patients with BHR implants at 5 years is lower than that of patients with Metasul metal-metal total hip replacements. The whole blood levels of cobalt and chromium are higher postoperatively than preoperatively, but there does not appear to be an increase in the levels over time. In addition, the whole blood levels of cobalt and chromium in very active individuals in the early postoperative period was not different than in usual patients. Based on a comparison of the measure Co and Cr levels with the recommended safe reference levels (EKA and BEI), the metal ion levels in patients with BHRs were in the safe range.

A long-term study of the cancer rates in 579 patients with historic metal-metal total hip replacements over a maximum period of 30 years was provided. There was no increase in either all-site cancer or site-specific cancer rates (Visuri T, Pukkala E. Does metal-on-metal hip prosthesis have influence on cancer? A long-term follow-up study. Eds. Reiker C, Oberholzer S, Wyss U. World Tribology Forum in Arthroplasty (pub), Hans Huber Bern, Toronto, Seattle: pp.181-188, 2001).

- A summary of the literature pertaining to the medium and long-term safety of cobalt and chromium ion exposure was provided and included copies of the following references:

Jacobs JJ, et al.: Cobalt and chromium concentrations in patients with metal on metal total hip replacements. Clin. Orthop., 329 (supplement): S256-S263, 1996. Abstract.

The authors measured the serum and urine concentrations of Co and Cr in 8 patients implanted with the McKee-Farrar metal-metal total hip replacements at greater than 20 years. There was a 9-fold elevation in serum Cr, 35-fold increase in urinary Cr, and 3-fold increase in serum Co. In 6 patients with metal-metal surface replacements, there was a 3-fold increase in serum Cr, 4-fold increase in urinary Cr, and a 4-fold increase in serum Co at less than 2 years.

Jacobs JJ, et al.: Metal release in patients who have had a primary total hip arthroplasty. A prospective, controlled, longitudinal study. J. Bone Joint Surg., 80(10): 1447-1458, 1998. Abstract.

The authors measured the serum and urine concentrations of Ti, Al, Co, and Cr in patients with metal-poly total hip replacements. At 36 months, patients had as much as a 3-fold increase in serum Ti levels if titanium implants were used. In patients with Co alloy implants, there was as much as a 5-fold and 8-fold increase in the concentrations of Cr in serum and urine, respectively. The modular head-neck junction was identified as a likely source of ion release.

Schaffer AW, et al.: Increased blood cobalt and chromium after total hip replacement. J. Toxicol. Clin. Toxicol., 37(7): 839-844, 1999. Abstract.

The authors found that there were significant postoperative elevations in urine Co and Cr and blood Co levels in all 76 patients, and in 29 patients the levels exceeded the EKA threshold limits for safe blood and urine Co levels.

Savarino, et al.: Ion release in stable hip arthroplasties using metal-on-metal articulating surfaces: a comparison between short- and long-term results. J. Biomed. Res., 66A(3): 450-456, 2003. Abstract.

The authors found that the serum Co and Cr levels were increased at 24 months and at 52 months. Delaunay CP found that there was no correlation between systemic Co concentrations and age, gender or patient activity. Ladon D found an increased incidence of chromosome translocations and aneuploidy in patients with both metal-metal and metal-poly total hip replacements.

Masse A, et al.: Ion release and chromosomal damage from total hip prostheses with metal-on-metal articulation. J. Biomed. Mater. Res. B. Appl. Biomater., 67(2): 750-757, 2003. Abstract.

The authors measured the Co, Cr, Ni and Mb levels in blood and urine after Metasul total hip replacements. The levels increased 2-fold (blood Co), 10-fold (urine Co), 1.5-fold (blood Cr), and 3-fold (urine Cr) at 6 months. There were no changes in the frequency of markers of chromosomal damage in the peripheral lymphocytes at any observation time points.

Visuri T, et al.: Cancer risk after metal on metal and polyethylene on metal total hip arthroplasty. Clin. Orthop., 329 Supplement: S280-S289, 1996. Abstract.

The authors state that the risk of total cancer in patients with a metal-metal McKee-Farrar hip replacement is 1.23-fold compared to metal-poly hip replacements at 15.7 years.

Visuri T, Pukkala E. Does metal-on-metal hip prosthesis have influence on cancer? A long-term follow-up study. Eds. Reiker C, Oberholzer S, Wyss U. World Tribology Forum in Arthroplasty (pub), Hans Huber Bern, Toronto, Seattle: pp.181-188, 2001.

The authors surveyed 579 patients who received a McKee-Farrar metal-metal hip replacement and had long-term follow-up (average 16.8 years). The annual incidence of all-site cancers was the same as expected. There was an excess of cancers with unknown primary site in women, a borderline excess of colon cancer after 15 years, higher number of leukemias, but a decreased number of urinary tract cancers. No bone or connective tissue sarcomas were observed. Other forms of cancer were the same as in the general population.

MacDonald SJ, et al.: Metal-on-metal versus polyethylene in hip arthroplasty: a randomized clinical trial. Clin. Orthop., 406: 282-296, 2003. Abstract.

Erythrocyte and urine metal ion levels were measured in 23 metal-metal and 18 metal-poly total hip replacement patients. 41% of the metal-metal patients had increasing metal ion levels at the latest follow-up. Patients with metal-metal THR had a 7.9-fold increase in erythrocyte Co, 2.3-fold increase in erythrocyte Cr, 35.1-fold increase in urinary Co, and a 17.4-fold increase in urinary Cr.

Maezawa K, et al.: Cobalt and chromium concentrations in patients with metal-on-metal and other cementless total hip arthroplasty. Arch. Orthop. Trauma Surg., 122(5): 283-287, 2002. Abstract.

The serum and urine concentrations of Co and Cr in 32 patients with metal-metal total hip replacements were measured and compared with 43 patients with metal-poly total hip replacements. The serum and urine Co concentrations were not detectable in any patients. The serum and urine Cr concentrations were elevated in 37.5% and 90.6% of metal-metal patients.

Witzlieb, WC, et al.: Histopatholgical findings and metal ion concentrations in Metasul and Birmingham Hip Resurfacing metal on metal bearings. U Hanisch, V Neumeister & WC Witzler, University of Dresden, Germany, June 2002.

The authors presented the results of 163 Birmingham Resurfacing Hip cases (number of Metasul cases not identified), including histopathology of 9 cases (5 BHR and 4 Metasul) and serum Co and Cr concentrations in 67 BHR and 32 Metasul patients (average 6 and 14 months, respectively). There were wear particles in only 2 of the 5 BHR cases. There was regular but not high amounts of metal debris in the Metasul patients. There were no inflammatory changes, foreign body reactions or metallic debris in the BHR capsular tissue. Both devices produced detectable serum Co and Cr levels by 1 month postoperatively, but these levels did not change over the course of the 44 months follow-up time. There were no significant differences in the serum ion levels between the BHR and Metasul patients.

Clarke MT, et al.: Levels of metal ions after small- and large-diameter metal-on-metal hip arthroplasty. J. Bone Joint Surg., 85B(6): 913-917, 2003.

The serum levels of Co and Cr were measured in 22 patients with metal-metal resurfacing and 22 patients with metal-metal total hip replacements. At 16 months, the median serum levels of Co and Cr were 38nmol/l and 53nmol/l, respectively, for the resurfacing patients and 22nmol/l and 19nmol/l, respectively, for the total hip patients.

Brodner W, et al.: Serum cobalt levels after metal-on-metal total hip arthroplasty. J. Bone Joint Surg., 85A(11): 2168-2173, 2003. Abstract.

The authors reported on the results of 50 Metasul metal-metal and 50 ceramic-poly total hip replacements. At 1 year, the median concentration of whole blood cobalt was 1.0μg/l and 0.7μg/l at 5 years in the metal-metal group, and undetectable in the ceramic-poly group.

Migaud H., et al.: Cementless metal-on-metal hip arthroplasty in patients less than 50 years of age. J. Arthroplasty, 19(8) Supplement 3: 23-28, 2004.

The authors reported on the results of 39 metal-metal total hip replacements. At a minimum of 5 years, the median concentration of whole blood cobalt was 0.62μg/l (range 0.2-4.7μg/l). Three women delivered healthy babies.

Delaunay, CP: Metal-on-metal bearings in cementless primary total hip arthroplasty. J. Arthroplasty, 19(8) Supplement 3: 35-40, 2004.

The authors measured the whole blood concentrations of cobalt in 99 patients who had Metasul metal-metal total hip arthroplasties out to 9 years. There were 76 patients with elevated postoperative Co levels (60 were in the laboratory "normal" range) and 23 patients that had unchanged levels.

Ladon D, et al.: Changes in metal levels and chromosome aberrations in the peripheral blood of patients after metal-on-metal hip arthroplasty. J. Arthroplasty, 19(8) Supplement 3: 78-83, 2004.

The authors found that there is a significant increase in the chromosome translocations and aneuploidy in lymphocytes at 6 months, 12 months and 24 months in patients with metal-on-metal hip arthroplasties who have elevated cobalt and chromium levels.

Jacobs, J, et al.: Can metal levels be used to monitor metal-on-metal hip arthroplasties? J. Arthroplasty, 19(8) Supplement 3: 59-65, 2004.

This is a review of the current practices of performing tests for metal ion concentrations in blood, serum, and urine in patients who have metal-on-metal hip replacements. The authors conclude that these tests are valuable research tools, but are not useful clinically to monitor patients for metal-related toxicity.

MacDonald SJ: Can a safe level for metal ions in patients with metal-on-metal total hip arthroplasties be determined? J. Arthroplasty, 19(8) Supplement 3: 71-77, 2004.

This paper is a review of previously reported studies of cobalt levels in total hip replacement patients and a discussion about the safety standards for metal ions. The author concludes that in order to determine whether there is a causal relationship between metal-on-metal bearings and any potential risk will require a significant number of patients.

Summary: Metal Ion Literature Analysis

These publications demonstrate that serum and urinary metal ion concentrations in patients with total hip replacement in general, and metal-metal implants in particular, increase in the postoperative period. However, there does not appear to be any conclusive evidence that elevated cobalt and chromium levels have any significant detrimental effects in total hip arthroplasty patients.

EFFECTIVENESS DATA

Survivorship

The survivorship estimates were based on the number of patients with no revision. Survivorship analyses were provided for various cohorts and demographic subgroups calculated according to Peto's adjustment method as follows (Table 19):

Table 19: % Survivorship Analyses (no revision)

Population	1 year	2 years	3 years	4 years	5 years
X-ray Cohort	99.2	99.2	98.4	98.4	98.4
Oswestry Cohort	99.4	99.0	98.7	98.6	98.4
X-ray/Oswestry Combined Cohort	99.4	99.0	98.7	98.6	98.4 (95% CI, 97.3-99.5%)
McMinn Cohort	99.6	99.6	99.6	99.6	99.6
Overall McMinn Cohort	99.4	99.1	98.8	98.7	98.5 (95% CI, 97.4-99.6%)
Male[1]	99.4	99.2	98.9	98.9	98.6
Female[1]	99.4	99.0	98.5	98.2	98.2
Age ≤65 years[1]	99.5	99.2	98.8	98.7	98.5
Age >65 years[1]	99.0	99.0	99.0	99.0	99.0
Dx: AVN[1]	98.9	98.9	96.7	96.7	92.1 (95% CI, 82.2-100%)
Dx: Dysplasia[1]	99.4	99.4	98.9	98.1	98.1
Dx: OA	99.5	99.1	98.8	98.8	98.8 (95% CI, 98.3-99.4%)

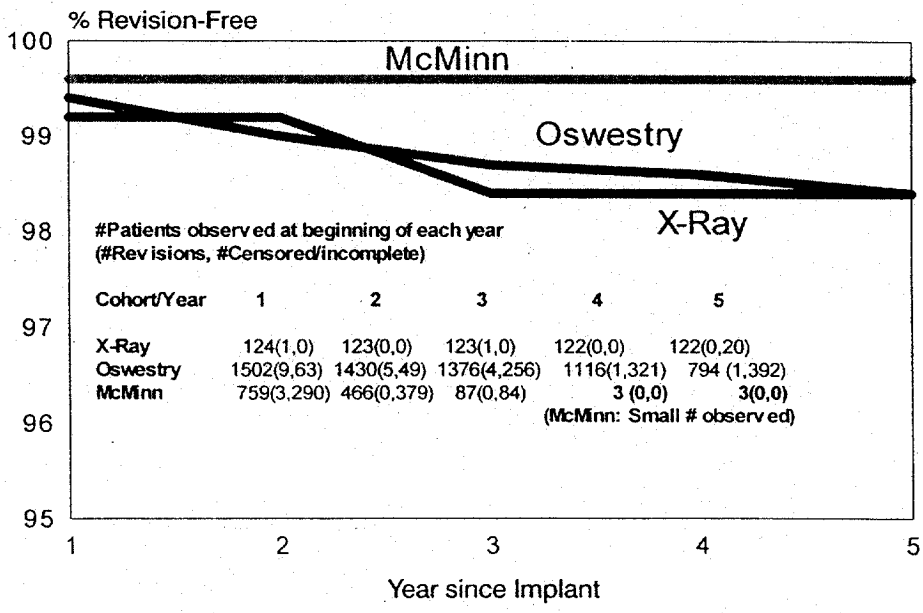
Table 19 Continued: % Survivorship Analyses (no revision)					
Population	**1 year**	**2 years**	**3 years**	**4 years**	**5 years**
Dx: Inflammatory[1]	98.1	98.1	98.1	98.1	98.1
Dx: Other[1]	100.0	100.0	100.0	100.0	100.0
Unilateral[1]	99.4	99.1	98.8	98.6	98.4
Bilateral[1]	99.6	99.2	98.8	98.8	98.8
Baseline OSHIP ≤63[2]	99.0	98.7	98.7	98.7	98.7
Baseline OSHIP >63[2]	99.8	99.3	98.7	98.3	98.3
Baseline OSHIP missing[2]	99.5	99.5	98.8	98.8	98.3
BMI ≤26[2]	99.7	99.3	99.0	98.8	98.8
BMI >26[2]	99.1	98.9	98.7	98.7	98.3
BMI missing[2]	99.4	99.1	98.1	98.1	98.1

[1] For the Overall McMinn cohort (2,385 hips)

[2] For the X-Ray + Oswestry cohorts (1,626 hips)

There were no statistically significant differences in cumulative 5-year survival (revision-free) probabilities among three study cohorts. The following Figure 1 summarizes these cumulative survival probabilities (all hips):

Figure 1. Cumulative % Revision-Free, BHR



Due to small number of revisions (total 25, ≤ 5-year follow-up) from large numbers in three study cohorts (total of 2385 hips), there were no statistically significant differences for all pairwise comparisons in 5-year survival (revision-free) probabilities among three cohorts, either by log-rank test, Wilcoxon test, or Cox proportional hazard (PH) regression analysis. Both the Cox PH regression model and the log-rank test require that the two survival probability curves be parallel or nearly parallel (no significant cohort by time crossover).

The above three statistical significance tests were also applied to several clinically important patient covariates, which include age (≤65, >65), gender (M, F), reason for resurfacing (AVN, osteoarthritis (OA), inflammatory arthritis (IA), dysplasia, and others; reference group = OA), baseline OSHIP score (yes, no), hips (unilateral, bilateral). The only marginally statistically significant difference in 5-year survival probability was between the patients with Osteoarthritis (98.8%) and Avascular Necrosis (92.1%) as their primary diagnostic indication. The p-values to compare these two % revision-free curves for OA versus AVN comparison are p=0.0415 (Log-rank) and p=0.2282 (Wilcoxon).

Due to non-parallelism of the Oswestry and X-Ray survival curves, careful clinical interpretation is needed. Both log-rank and Wilcoxon test that the two revision-free curves are equal, and the Cox PH model tests that the ratio of the two hazards (probability of revision) is unity. The log-rank test assigns *equal weight* to *all* follow-up times and the Wilcoxon test assigns *more weight* to the *earlier* follow-up times where more patients are at risk of revision. The log-rank test has optimum statistical power if the parallelism assumption for the two revision-free curves is valid. The Cox PH model is not appropriate here due to obvious non-parallelism of the two curves in Figure 1. The percentages of revisions are 3.1% (3/97) for AVN, 1.1% for dysplasia (4/377), 0.95% (17/1789) for OA, 1.7% (1/57) for Inflammatory arthritis (IA), and 0% for others (0/65), with a combined 1% (25/2385) revisions over all diagnostic groups, during 5-year follow-up.

There were 37 cases (of the 1626 cases) with a diagnosis of "Other." There were no revisions in this group, and thus the survivorship at 5 years is 100%. A separate analysis for this "Other" group was not provided and approval for indications other than OA, IA, AVN and DDH was not proposed.

Radiographic Data

The clinical data used to support this series contained the results of an independent radiographic review of the X-Ray Cohort, the first 124 procedures performed in the series from July 1997 through December 1997.

Radiographs were taken on 108 of the 118 procedures expected at 5 years postoperatively (91.5%). Six (6) procedures were not expected at 5 years postoperatively because one patient with bilateral hip implants died from a motor neuron disease unrelated to the BHR procedure; and 4 of the 124 BHR procedures (3.2%) have undergone revision: 3 cases were revised for infection, and 1 case required revision because of a femoral neck fracture. Therefore, 118 procedures (124 hips - 2 hips due to death - 4 revisions = 118 procedures) were eligible for 5 year radiographic evaluation of the BHR. Ten other cases were missing due to lost to follow-up or incomplete film records. Therefore, one hundred and eight (108) of the 118 hips surviving to 5 years had 5 year radiographs available for independent review (91.5%). (Note: An additional bilateral patient died 7 years post-op due to stroke but had 5 year x-rays taken).

Baseline films for the purposes of comparisons were made in each of the 108 cases in the postoperative time period (usually within 3 months, but 8 of the 108 procedures had baseline evaluations performed at time points ranging from 110-860 days).

Radiographic Study: 5-Year Radiographic Assessments

The radiographs were assessed for radiolucencies, bone resorption, heterotopic bone, acetabular angle, medial-lateral migration, and other observations to determine whether a revision surgery was necessary.

Femoral radiolucencies: Radiolucencies were graded 0-9 (Amstutz scale). There were femoral radiolucencies found in 4 cases (4.1%)—1 each with grade 9 (migration), grade 5

(zone 2-3), grade 2 (zone 1) and grade 1 (zone 2). The patient with a grade 9 femoral radiolucency was classified as a radiographic failure.

Acetabular radiolucency: Radiolucencies were graded 0-9 (DeLee and Charnley scale). There were 2 hips with acetabular radiolucencies, both with grade 8 (zones I-III, complete) findings. One hip had preoperative acetabular cysts that progressed over time, and the other had a preoperative dysplastic acetabulum and developed protrusio. Both were classified as radiographic failures. Three patients had insignificant radiolucencies (grade 1 in two hips and grade 2 in one hip).

Heterotopic bone: There were 21 hips that had Brooker I and 5 hips with Brooker II heterotopic ossification (HO). Only 2 hips had "clinically significant HO," (i.e., Brooker III or IV). Both had Brooker III HO. Thus, 28 of the 108 procedures evaluated (28.9%) had any heterotopic bone at 5 years and 2.1% had significant HO. None of the cases with heterotopic bone were determined to require a revision.

Acetabular angle: There was only 1 case that had a change in the acetabular angle >5°. This patient also had the grade 8 acetabular radiolucency (see above). No cases had a change in acetabular angle that was determined to be an indication for a revision.

Medial / Lateral Migration: There were no procedures with a change in medial/lateral acetabular cup position, and no cases with a change in acetabular position that was determined to be an indication for a revision.

Additional observations: Bone resorption at the femoral neck was found in 3 cases. In no case was the resorption associated with any other notable radiographic findings. Bone cysts were found in 2 patients: one, described above, and the other had 3cm cysts associated with a grade 1 acetabular radiolucency. No other significant signs were noted.

Three (3) of the 108 (2.8%) patients for whom radiographs were available were radiographic failures at 5 years (Table 20).

Table 20: Radiographic Findings
Number of procedures (%)

Findings	Number (%)
Femoral radiolucencies	
Failure: Grade 9	1 (0.9%)
Other: Grade 1	1 (0.9%)
Other: Grade 2	1 (0.9%)
Other: Grade 5	1 (0.9%)
Acetabular radiolucencies	
Failure: Grade 8[1]	2 (1.8%)
Other: Grade 1	2 (1.8%)
Other: Grade 2	1 (0.9%)
Change in orientation/migration	
5° change in orientation[1]	1 (0.9%)
Heterotopic ossification	
Brooker IV	0 (0.0%)
Brooker III	2 (1.8%)
Brooker II	5 (4.6%)
Brooker I	21 (19.4%)
Other	
Bone resorption, femoral neck	3 (2.8%)
Femoral or acetabular cyst	2 (1.8%)

[1] Occurred in the same patient

Radiographic Study: Comparison to Literature Reference
The radiographic results were compared and found to be similar to the literature reference group (Table 21).

Table 21: Radiographic Findings
X-Ray Cohort vs. Literature Reference

Radiographic Finding	Overall McMinn Cohort	Garino Reference*	D'Antonio Reference		
			ABC with porous (n=162)**	ABC with HA (n=169)**	Reference Control M/PE (n=149)**
Femoral RL zone 1	1 (0.9%)	-	4 (2.5%)	4 (2.4%)	6 (4.0%)
Femoral RL zone 2	1 (0.9%)	-			
Femoral RL zone 2 & 3	1 (0.9%)	-			
Femoral RL zone 7	0	-	2 (1.2%)	1 (0.6%)	0
Stem subsidence	0	-	0	1[1] (0.6%)	0
Unstable stem	1 (0.9%)	-	0	1[1] (0.6%)	0
Cup RL Zone I	2 (1.8%)	-	10 (6.2%)	1 (0.6%)	10 (6.7%)
Cup RL Zone II	1 (0.9%)	-	3 (1.9%)	0	7 (4.7%)
Cup RL Zone III	0	-	25 (15.4%)	0	35 (23.5%)
Cup RL all 3 zones	2 (1.8%)	-	0	0	0
Cup migration	1 (0.9%)	-	0	0	1[2] (0.7%)
Cup unstable		-	1 (0.6%)	0	1[2] (0.7%)

* No radiographic data.
** Revision rates are based on a minimum of 2-year follow-up and available x-rays.
[1] Same femoral component
[2] Same acetabular component

Pain and Function - Oswestry Modified Harris Hip (OSHIP) Score—Unilateral Procedures Only
FDA believes that it is difficult to assess the pain and function of each hip separately in patients with bilateral hip involvement using the Harris Hip Score or the Oswestry-modified Harris Hip Score (OSHIP), because it is difficult to distinguish the contributions of each hip on functional assessments such as walking or support, walking distance, stair-climbing, sitting, and transportation. Therefore, FDA believes only the unilateral patients should be used in an analysis of pain and function for the purposes of evaluating safety and effectiveness.

The mean OSHIP Scores (unilateral procedures only) improved from a baseline mean of 60.1 to 94.8 at 5 years. For the group of patients who had high baseline OSHIP scores (≥80), the mean OSHIP scores improved from 84.5 to 99.3. The group of patients who had low baseline OSHIP scores (<80), the mean OSHIP scores also improved from 59.4 to 95.6. At postoperative years 2, 3, 4 and 5, the percentage of cases with good or excellent OSHIP scores was 96.9%, 95.8%, 95.2%, and 92.8%, respectively (Table 22).

Table 22: Oswestry-Modified Harris Hip Score (OSHIP) X-Ray / Oswestry Combined Cohort—Unilateral only						
	Baseline	**1 year**	**2 years**	**3 years**	**4 years**	**5 years**
Expected	1111	1103	1100	927	687	395
OHSIP assessments	892	835	842	818	607	360
OSHIP mean	60.1	96.6	96.8	96.2	95.9	94.8
SD*	13.1	6.75	7.3	7.4	8.0	9.7
SE**	0.44	0.23	0.25	0.26	0.32	0.51
95% CI	(59, 61)	(96, 97)	(96.3, 97.3)	(95.7, 96.9)	(95.2, 96.6)	(93.8, 95.8)
AVN OSHIP mean	49.4	91.3	93.6	96.2	94.3	97.4
N, AVN	43	35	38	32	23	14
Dysplasia OSHIP mean	57.7	96.2	96.7	95.2	94.7	90.6
N, Dysplasia	131	123	117	117	81	44
OA OSHIP mean	61.5	97.0	97.0	96.5	96.2	95.3
N, OA	678	642	652	632	484	287
IA OSHIP mean	48.5	95.5	94.9	93.2	91.6	89.3
N, IA	15	11	11	15	10	8
Other OSHIP mean	62.9	96.5	98.3	96.6	98.8	98.4
N, Other	25	24	24	22	9	7
OSHIP mean for procedures with baseline ≥80	84.5	96.1	97.8	97.3	99.6	99.3
N, for baseline ≥80	25	22	22	18	8	3
OSHIP mean for procedures with baseline <80	59.4	96.9	96.9	96.6	96.4	95.6
N, for baseline <80	867	693	686	635	440	240
OSHIP mean for procedures with baseline OSHIP	60.1	96.9	96.9	96.6	96.5	95.6
N, with baseline OSHIP	892	715	708	653	448	243
OSHIP mean for procedures without baseline OSHIP	-	94.8	96.2	94.8	94.1	92.9
N, without baseline OSHIP	-	120	134	165	159	117
Improved ≥10 (%)	-	703 (84.2)	697 (82.8)	645 (78.9)	445 (73.3)	239 (66.4)
Maintained (%)	-	130 (15.6)	142 (16.9)	173 (21.1)	161 (26.5)	121 (33.6)
Deteriorated ≥10 (%)	-	2 (0.2)	3 (0.4)	0	1 (0.2)	0
OSHIP Excel ≥90 (%)	2 (0.2)	757 (90.7)	775 (92.0)	722 (88.3)	529 (87.1)	307 (85.3)
OSHIP Good 80-89 (%)	23 (2.6)	56 (6.7)	41 (4.9)	61 (7.5)	49 (8.1)	27 (7.5)
OSHIP Fair 70-79 (%)	175 (19.6)	12 (1.4)	14 (1.7)	20 (2.4)	16 (2.6)	12 (3.3)
OSHIP Poor 60-69 (%)	349 (39.1)	3 (0.4)	5 (0.6)	9 (1.1)	8 (1.3)	8 (2.2)
OSHIP V Poor <60 (%)	343 (38.5)	7 (0.8)	7 (0.8)	6 (0.7)	5 (0.8)	6 (1.7)

*SD = Standard deviation; **SE = Standard error of sample mean = SD/$\sqrt{n}$; CI = confidence interval of true OSHIP mean

For the data in the table above regarding the number of procedures who improved ≥ 10 pts., maintained, or deteriorated ≥ 10 pts., it was explained that those patients with no baseline scores were counted as "maintained." The table below contains an analysis of the number of procedures who improved ≥ 10 pts., maintained, or deteriorated ≥ 10 pts., when the patients without baseline scores are removed from this analysis and just counted as missing (Table 23).

Table 23: OSHIP Improvement Oswestry and X-Ray Cohorts						
	Change	**1 year**	**2 years**	**3 years**	**4 years**	**5+ years**
Unilateral	Improve ≥10	703 (98.3)	697 (98.4)	645 (98.8)	445 (99.3)	239 (98.4)
	Same <10	10 (1.4)	8 (1.1)	8 (1.2)	2 (0.4)	4 (1.6)
	Worse ≥10	2 (0.3)	3 (0.4)	0 (0.0)	1 (0.2)	(0.0)
	N	715	708	653	448	243
	Missing	388	392	274	239	152

Pain and Function - Comparison to Literature References

In the literature references, the authors used Harris Hip Score, not OSHIP, to collect pain and function effectiveness data. D'Antonio *et al.* reported Harris Hip Scores at 2 - 4 year follow up (mean 3 year) for the ceramic-on-ceramic hip procedures as follows:

- ABC System 1 (porous): 95.4 mean score (n=166)
- ABC System 2 (HA): 96.6 mean score (n= 172)

Garino reported an average increase in Harris Hip Score from 44 pre-operatively to a mean of 97 at follow up. Although the pain and function data in the literature and in the BHR clinical data series were collected using different scoring systems, according to published and unpublished literature, patient self-assessments of hip pain and function including the OSHIP assessment method have been shown to produce lower scores; therefore provide a conservative estimate, as compared to the physician administered HHS assessment method.

Patient Satisfaction

The patient satisfaction question is not a standard component of the OSHIP assessment but was an additional question asked for this study in the annual, patient-completed, mail-in questionnaire. At 5 years, 99.5% of the procedures in the X-Ray/Oswestry combined cohort were pleased or very pleased with the operation. At 5 years, 99.2% of the unilateral procedures from the X-Ray/Oswestry combined cohort were pleased or very pleased with the operation (Table 24).

Table 24: Patient Satisfaction X-Ray/Oswestry Combined Cohort						
	X-Ray/Oswestry Combined Cohort N=1626					
	Base	**1 year**	**2 years**	**3 years**	**4 years**	**5+ years**
N	1626	1616	1607	1349	1007	601
Pleased	-	75 (6.1%)	62 (5.0%)	80 (6.7%)	50 (5.6%)	31 (5.7%)
Very pleased	-	1109 (89.6%)	1177 (94.7%)	1100 (92.7%)	839 (94.1%)	512 (93.8%)
X-Ray/Oswestry Combined Cohort - Unilateral Procedures Only						
# All Unilateral	1111	1103	1100	927	687	395
Assessments	892	835	842	818	607	360
Please/Very Pleased (VP)	-	800 (95.8%)	839 (99.6%)	813 (99.4%)	604 (99.5%)	357 (99.2%)
N, AVN	43	35	38	32	23	14
AVN Please/VP	-	35 (100.0%)	38 (100.0%)	32 (100.0%)	23 (100.0%)	14 (100.0%)

Table 24 Continued: Patient Satisfaction X-Ray/Oswestry Combined Cohort						
	X-Ray/Oswestry Combined Cohort N=1626					
	Base	1 year	2 years	3 years	4 years	5+ years
N, Dysplasia	131	123	117	117	81	44
Dysplasia Please/VP	-	119 (96.8%)	117 (100.0%)	115 (98.3%)	80 (98.7%)	43 (97.7%)
N, OA	678	642	652	632	484	287
OA Please/VP	-	613 (95.5%)	649 (99.6%)	630 (99.7%)	482 (99.6%)	285 (99.3%)
N, IA	15	11	11	15	10	8
IA Please/VP	-	11 (100.0%)	11 (100.0%)	15 (100.0%)	10 (100.0%)	8 (100.0%)
N, Other	25	24	24	22	9	7
Other Please/VP	-	22 (91.7%)	24 (100.0%)	21 (95.5%)	9 (100.0%)	7 (100.0%)

APPLICABILITY OF THE FOREIGN DATA FROM A SINGLE INVESTIGATOR AND UNITED KINGDOM PRACTICE OF MEDICINE TO THE TARGET UNITED STATES POPULATION AND PRACTICE OF MEDICINE

Comparison of the United States and United Kingdom Patient Populations

The clinical data series was derived from a foreign clinical study conducted by a single investigator at the Birmingham Nuffield and Little Aston Hospitals in the United Kingdom. There are no racial or ethnic origin data for the patients in the clinical data series. However, the racial and ethnic distributions in the U.S. and U.K. populations are similar. There were noted differences in the higher percentage of people of African-descent and "other races" in the general US population as compared to the general U.K. population (Table 25). However, in addition to the comparison of the U.S. and U.K. populations, additional unpublished data was provided on 3,374 BHR hips implanted by 140 surgeons and published reports from the experience of multiple surgeons implanting over 3,800 BHR hips in several countries around the world to support the applicability of data to the U.S. population and medical practice.

Table 25: Comparison of the Ethnic / Racial Distributions in General US/UK Populations		
	U.S.	U.K.
White	75.1%	92.1%
Black	12.3%	2.0%
Asian	3.6%	4.0%
Native American	0.9%	-
Pacific Islander	0.1%	-
Chinese	-	0.4%
Other race	5.5%	0.4%
Mixed race	2.4%	1.2%

A comparison of the demographics and diagnostic indications for the BHR study and a literature reference by D'Antonio and co-workers regarding the Howmedica Osteonics ABC/Trident Ceramic/Ceramic hip system was provided. There are noted differences in the higher percentage of men, higher percentages of procedures with dysplasia and inflammatory diagnostic indications and lower percentage of procedures with post-traumatic arthritis for the BHR study as compared to the ceramic/ceramic group (Table 26). However, there were no statistically significant differences in gender or in patients with diagnostic indications of dysplasia, inflammatory

arthritis, and post-traumatic arthritis as compared to osteoarthritis in their 5-year survival probability in the clinical data series. Another noted difference was a lower percentage of procedures with AVN for the BHR study as compared to the ceramic/ceramic group (Table 26). Patients with AVN were marginally statistically significant different in 5-year survival probability as compared to patients with osteoarthritis in the clinical data series, 98.8% and 92.1% respectively. However, the product labeling, operative technique, and training of user surgeons on use of the device should help to minimize the impact of this difference in the U.S. patient population.

Table 26: Patient Demographics and Diagnostic Indication Comparisons

	Overall BHR Metal/Metal Resurfacing Hip System	***D'Antonio*: HowOst C/C THR***
Hips	2385	514
% Men (n)	70.6% (1683)	65%
% Women (n)	29.4% (702)	35%
Mean Age (range)	53.1 (13.4-86.5)	53
% Age ≤65 years (n)	91.9% (2191)	-
Dx: % OA (n)	75% (1789)	78%
Dx: % DDH (n)	15.8% (377)	-
Dx: % AVN (n)	4.1% (97)	16%
Dx: % Inflammatory (n)	2.4% (57)	-
Dx. % Other (n)	2.7% (65)	2%
Dx: % Post-Traumatic Arthritis (n)	-	4%

* Data presented by the applicant taken from D'Antonio, J., Capello, W., Manley, et al., "New experience with alumina-on-alumina ceramic bearings for total hip arthroplasty," J. Arthroplasty 17(4): 390-97, 2002.

The applicability of the foreign data to the US patient population is based on its large sample size, as well as the comparable demographics and diagnostic indications to the multi-center literature reference group.

Description of the Single Investigator's Practice of Medicine

All of the surgeries on the 2,385 cases in this PMA were performed by a single investigator (with assistance from other surgeons) at the Birmingham Nuffield Hospital (except for 6 cases that were performed at the Little Aston Hospital, Birmingham, U.K.). The practice of medicine, specifically the orthopedic practice of medicine, utilized by the investigator is considered to be similar to the standard of orthopedic practice in the U.S. The investigator's standard peri-operative regimen was described as follows:

- Laminar air flow operating rooms with body exhaust suits
- Posterior surgical approach
- Standard surgical technique (described in the Surgical Technique Manual)
- Antibiotic prophylaxis intraoperatively and for 24 hours postoperatively (1.5g Cefuroxime)
- DVT prophylaxis using a single-dose (800 IU) intravenous heparin intraoperatively and compression stockings and low-dose aspirin postoperatively for 6 weeks
- Intraoperative venting of the femoral shaft to prevent fat/marrow emboli
- Early ambulation: full weight-bearing with a walker on postoperative day #1, progressing to crutches and canes
- Hospital discharge at postoperative day #6
- After 6 weeks postoperatively, begin range of motion exercises
- Recommended activities include swimming, pool exercise, non-impact or low-impact exercise at a gym; and, avoidance of high impact exercises during the first postoperative year

ADDITIONAL DATA SOURCES

The main data sources were presented above but additional, less complete data on 3,374 BHR cases performed by 140 surgeons worldwide (other than the single investigator) was summarized. This is called the **Worldwide/Other Cohort**.

Demographic information for the Worldwide/Other Cohort included gender, age, diagnosis, BMI, baseline OSHIP scores. The study cohort demography was similar in the Worldwide/Other Cohort and the X-Ray/Oswestry combined cohort, with the mean age of 53.0 years in the X-Ray/Oswestry combined cohort and 52.5 years in the Worldwide/Other Cohort. The diagnostic indications were somewhat different between cohorts: OA (78% X-Ray/Oswestry combined cohort vs. 90.8% Worldwide/Other Cohort).

A comparison of the revisions and survivorship estimates for the X-ray/Oswestry combined cohort versus the Worldwide/Other Cohort was provided. The primary reason for revision in the Worldwide/Other Cohort was a fracture in 34 cases (1.0%), loosening in 26 cases (0.8%), infection in 7 cases, AVN in 5 cases, dislocation in 5 cases, miscellaneous device failures in 5 cases, pain in 3 cases, and unknown in 3 cases (Table 27).

Table 27: Revisions

	Preop	1 year	2 years	3 years	4 years	5+ years
X-Ray/Oswestry Combined Cohort N=1626						
Number of procedures*	1626	1626	1553	1499	1238	916
Revisions	-	10	5	5	1	3
Survivorship estimates	-	99.4	99.0	98.7	98.6	98.4
Worldwide/Other Cohort N=3374						
Number of procedures*	3374	3374	3051	2888	2493	1417
Revisions	-	35	15	14	7	5
Survivorship estimates	-	98.7	98.0	97.5	97.0	96.3

* The number of procedures is the number of hips that were surviving at the end of the previous year based on the survival analysis. Note that for the Survivorship data the "year 1" data is starting from day 1 and the "year 2" data is starting from day 366, etc.

The Worldwide/Other Cohort patients had slightly lower OSHIP scores at all time points (Table 28).

Table 28: OSHIP Worldwide/Other Cohort

	Baseline	1 years	2 years	3 years	4 years	5 years
Worldwide OSHIP assessments	395	2356	2492	2364	1379	505
Worldwide Mean OSHIP	56.95	91.67	92.47	92.45	91.86	89.77

XI. CONCLUSIONS DRAWN FROM THE STUDY

FDA believes that the applicant has provided an adequate device description and the pre-clinical testing information provided a reasonable assurance of device safety. FDA believes that the safety data collection methods used in this study; including, the OOC procedures for collecting safety data with the annual mail-in questionnaire; Mr. McMinn's follow-up procedures including his data collected at the McMinn Center and in consultation with primary care physicians; and the metal ion literature analysis were reliable and provide for valid scientific evidence as defined in 21 CFR 860.7. FDA believes that the effectiveness data collection methods are adequate. In reviewing the information provided by the applicant to address how the 5-year OSHIP data was collected, how the OSHIP scoring system was developed, and the justification for its use, the OSHIP data in conjunction with the 5-year survivorship, radiographic, and patient satisfaction data are reliable and accurate and provide for valid scientific evidence as defined in 21 CFR 860.7. FDA believes that because the large number of BHR procedures in the clinical data series evaluated at the 2-year (1,243 procedures) and 5-year (546 procedures) follow-up timepoints; the additional unpublished data on 3,374 BHR hips implanted by 140 surgeons; and published reports from the experience of multiple surgeons implanting over 3,800 BHR hips around the world, the applicant has adequately addressed how the clinical data is applicable to the U.S. population and medical practice, is sufficient to demonstrate the safety and effectiveness of the device, and to ensure reproducibility of clinical results. Therefore, FDA believes it is reasonable to conclude that the benefits of the use of the BHR System for the target population outweigh the risk of illness or injury when used in accordance with the directions for use.

XII. PANEL RECOMMENDATION

The Orthopaedic and Rehabilitation Devices Panel (the Panel) met on September 8, 2005 in Gaithersburg, MD to make a recommendation to the FDA on the approvability of the Smith and Nephew, Inc. Birmingham Hip Resurfacing (BHR) System, P040033. FDA received expert clinical opinion from the Panel regarding the safety and effectiveness data collection methods, the applicability of the foreign data from a single investigator and United Kingdom practice of medicine to the target United States population and practice of medicine, and the study results with respect to the device's safety and effectiveness.

Regarding the safety data collection methods, some Panel members expressed concerns about what they considered a lack of prospectively collected information and that the safety information was coming from one source. The applicant clarified that the OOC collected data prospectively starting in 1997 and an independent group compiled all safety data for the case series retrospectively. Other Panel members indicated that because of the large series and the corresponding amount of safety data included, there was enough information to evaluate device safety. Also, Panel members stated that the methods used met the FDA definition of valid scientific evidence in 21 CFR 860.7.

Regarding the effectiveness data collection methods, some Panel members supported the use of patient questionnaires to capture safety and effectiveness data stating that it may reduce bias associated with patients wanting to please their physicians in their responses to physician administered questionnaires. Other Panel members stated that although patient questionnaires are

important, physical exams and radiographic data is also as important; these Panel members would have used a more conventional evaluation (i.e., HHS) rather than OSHIP.

Regarding whether or not the clinical data is applicable to the target US population, practice of medicine, and US orthopedic surgeon population, Panel members expressed concerns that the case series did not contain data on the variability of the use of the device at various centers or that the data would provide reassurance of its applicability to the US population at risk. Other Panel members expressed support for the clinical data stating: (1) that if a randomized, controlled, trial would have been performed in the US for this device, the applicant would have recruited a small group of 5 excellent hip surgeons who would be very different from the average hip surgeon; therefore, the only difference is that data has one learning curve rather than 5; (2) that the UK population appeared very similar to the US population including the use of referral practices; (3) that there was additional literature information on the use of the device by other surgeons; and (4) that with an adequate training plan and proposed post-approval study these concerns may be mitigated.

The Panel voted three to two to recommend that FDA approve the PMA with conditions. The recommended condition of approval was as follows: The applicant should conduct the proposed post-approval study presented in the PMA with the addition of a clinical and radiographic evaluation at the 10-year follow-up time point. In addition, the sample size for the post-approval study should be based on statistical principles and the criteria for success.

XIII. CDRH DECISION

The PMA was filed on July 19, 2004 and granted expedited review status. The BHR was granted expedited review status because total hip systems with a resurfacing femoral component and a metal-on-metal articulation may offer advantages in safety and effectiveness over existing alternatives; such as, the preservation of femoral bone stock during implantation as compared to metal-on-metal total hip systems and a decrease in adverse tissue reaction due to particulate wear debris as compared to metal-on-polyethylene resurfacing hip systems.

CDRH agreed with the Panel's recommendation for the BHR System (approvable with conditions). The applicant has adequately submitted all information requested by CDRH for their Premarket Approval application.

The applicant's manufacturing facilities were inspected and were found to be in compliance with the Quality System Regulation (21 CFR 820).

CDRH has determined that because the large number of BHR procedures in the clinical data series evaluated at the 2-year (1,243 procedures) and 5-year (546 procedures) follow-up timepoints; the additional unpublished data on 3,374 BHR hips implanted by 140 surgeons; and published reports from the experience of multiple surgeons implanting over 3,800 BHR hips around the world, the use of this device for the labeled indications has been shown to be reasonably safe and effective.

CDRH believes that in addition to the post-approval requirements outlined in the approval order enclosure, the applicant must provide the following data every 6 months for the first two years and annually thereafter following PMA approval until completion of the post-approval studies and the submission of a final report:

1. The applicant has agreed to conduct a study to evaluate longer-term safety and effectiveness of the Birmingham Hip Replacement (BHR) System. This study is expected to include the first 350 consecutive cases of the 2,385 cases in the Overall McMinn Cohort that were included in the PMA. Patient pain, function, movement, revision status, and adverse events will be assessed at baseline and annually from five (5) years through ten (10) years post-op through the use of the Oswestry-Modified Harris Hip (OSHIP) patient self-assessment questionnaire. In addition, a clinical and radiographic examination will occur at ten (10) years post-op to evaluate adverse events, revisions, and evidence of any radiolucencies, osteolysis, or component position change.

2. The applicant has agreed to conduct a study to evaluate the learning curve, training program, and longer-term safety and effectiveness of the BHR System in the United States. This study will assess the generalization of the experience from a single physician in the United Kingdom to medical practice in the United States. Results of this study will be reflected in the labeling.

   This study is expected to include 350 patients at up to 8 sites with a minimum of 35 patients per site. Investigational sites recruited to participate in this study will be comprised of a geographically diverse mix of academic, referral, and/or community based sites. Per the training program outlined in the PMA, investigators will be recruited from the "Core Surgeon" group who will initially be trained on the BHR System as well as from a group of other interested US Surgeons who are subsequently trained by the Core Surgeons.

   Clinical and radiographic data will be assessed at baseline, annually through five (5) years post-op, and at ten (10) years post-op. Pain, function, and range of motion evaluations will be performed using the Harris Hip Score assessment method. Radiographs will be collected to evaluate any radiolucencies, osteolysis, or component position change. In addition, revision and adverse event data will be collected.

   The subjects will be assessed in the interim years of six (6) through nine (9) years post-op by use of a "postcard" (mailed, e-mailed, or telephone call) questionnaire follow-up to assess each subject's general well-being, and to determine if the study components remain implanted or are revised.

   In addition, to further assess the safety of the BHR system, cobalt and chromium ion concentration in the blood and renal function data (such as creatinine, GFR, BUN), will be collected preoperatively and at the 1-, 4, and 10-year follow-up timepoints.

3. The applicant has agreed to implement a training program, as outlined in the PMA. The training program includes quarterly investigator teleconferences or meetings for the first two years of the US study to provide a clinical update to investigators; to discuss study issues including adverse events; and to identify recommendations for improvement of the training program or labeling. If some investigators cannot attend the conference, the applicant has agreed that these investigators will all be contacted by telephone or will be sent the "Investigator Feedback Form" so that individual feedback can be obtained. The applicant has agreed to submit a summary of the minutes of the quarterly teleconferences/ physical meeting/ investigator feedback information as part of the report.

4. The applicant has agreed to provide an analysis of adverse events and complaints (including MDRs) received regarding the BHR system. In addition, the applicant agreed to use this analysis to provide a justification for modifications to the training program, post-approval study, labeling, and/or device design. Any modification to the post-approval study, labeling, and/or device design will be submitted for FDA review and approval prior to implementation.

The applicant was advised that the results of the post-approval studies, training program assessment, and adverse event analysis outlined in items 1-4 above must be reflected in the labeling (via a supplement) when the post-approval study is completed, and/or at earlier timepoints, as needed. The applicant agreed to this post-approval condition.

FDA issued an approval order on May 9, 2006.

XIV. APPROVAL SPECIFICATIONS

Directions for use: See the labeling.

Hazards to Health from Use of the Device: See Indications, Contraindications, Warnings and Precautions, and Adverse Events in the labeling.

Postapproval Requirements and Restrictions: See approval order.