

Division of General and Restorative Devices  
Orthopedic Devices Branch, HFZ-410

Food and Drug Administration  
Office of Device Evaluation  
9200 Corporate Boulevard  
Rockville, MD 20850

**Date:** April 19, 2004  
**To:** File  
**From:** Medical Officer  
**Subject:** Hip Guidance Document Submission:  
Clinical Trial Design for Hip replacement Systems  
**Sponsor:** Orthopedic Surgical Manufacturers Association  
**Contact:** Joel Batts 813-877-4469  
**Docket #**

### **Background**

Once a year, FDA publishes, both in the **Federal Register** and on the Internet, a list of possible topics for future guidance document development or revision during the next year according to 21 CFR IB-10.115(5). This document is the Orthopedic Surgical Manufacturers Association's (OSMA's) response to this list and discussion between industry, and professional societies about the needs for assisting in getting new devices to the public in accordance to the MDUFMA goals. In accordance with 21 CFR IB-10.115(3), anyone can submit drafts of proposed guidance documents for FDA to consider. These documents are termed "Guidance Document Submissions"(GDS) and are submitted it to Dockets Management Branch of FDA. FDA is required to acknowledge receipt of the "Clinical Hip" GDS and provide public access to it on the Dockets web site, per the FDA's Good Guidance Practices (GGPs) at 21 CFR 10.115.

The subject of this GDS is a clinical study design for evaluating hip joint replacement systems, an all encompassing term, to include both conventional hip arthroplasty prostheses (cemented and noncemented) and modern technological improvements on the conventional designs. It contains three objective performance criteria as a composite patient success criterion and a performance benchmark for study success.

This is the Office of Device Evaluation's first GDS which is one of the reasons for presenting it to the advisory panel. There are many examples of objective performance criteria used for preclinical data comparisons. In addition, there exist examples where acceptable endpoint criteria based on meta-analysis of medical literature are used as a control against which the safety and effectiveness of emerging technological devices are measured. Examples include cardiac heart valve devices, cardiac ablation catheters and intraocular ophthalmic lens devices.

It should be emphasized that this is *not* a petition for reclassification of the hip replacement systems. The devices that may be affected include those described as classified devices by 21 CFR 888.3310, 888.3320, 888.3330, 888.3340, 888.3350, 888.3353, 888.3358 and 888.3410 and potential unclassified new types of devices. This information is not intended to supplant current guidances which define standard preclinical investigations that characterize hip replacement prostheses. Any subsequent guidance that develops would be used to supplement those guidances.

### **The Premise**

The ideal and preferred vehicle for establishing safety and effectiveness for any device is the randomized clinical trial. However, FDA's regulations implementing section 513(a)(3) of the Act establish a hierarchy of valid scientific evidence. Under 21 CFR 860.7(c)(2), endorsed by the US Preventive Services Task force Quality of Evidence (USPSTF-HHS) includes

- well-controlled investigations,
- partially controlled studies,
- studies and objective trials without matched controls,
- well documented case histories, 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 the safety and effectiveness of a device under its condition of use.

Depending on the availability of data from previous studies of these devices and previously studied similar devices in comparable populations, other sources of valid scientific evidence to show safety and effectiveness on this list can be considered in an effort to provide a less burdensome approach to premarket applications. Such an approach is one of the intentions of this document. It proposes to provide a standardized method for designing clinical trials intended to evaluate the safety and effectiveness of hip joint replacement systems, with minimum benchmark achievement goals for patient and study success, in an effort to provide a least burdensome approach to manufacturers and facilitate FDA review. A fundamental concept underlying "least burdensome" concept is that safety and effectiveness be demonstrated by appropriate and valid scientific information, evidence, or data. Least burdensome is not intended as a way for either the FDA or Industry to "cut corners" regarding the generation of data to support a marketing application.<sup>3</sup>

The underlying rationale for this document is a long history of knowledge and understanding of conventional hip joint arthroplasty dating to before the 1960s. Currently, there is considerable cumulative experience on prosthetic hip joint clinical performance, given the 40 years of clinical evolution and an implantation rate of approximately 300,000 or more devices/year.<sup>7</sup> Over the past 10-20 years total hip arthroplasty has become the standard of care for end stage arthritic and other medical conditions affecting the weight bearing surfaces of the hip joint. Improved quality of life, through the reduction of pain and return to function are well accepted outcomes in the professional, scientific and patient communities. There is an abundance of literature with long-term follow-up available that documents the safety and effectiveness for conventional hip arthroplasty which allows comprehensive review and understanding by the industry, physicians and FDA. This understanding allows for the development of benchmark criteria, by which to measure new emerging technologies in the least burdensome way.

### **How can the FDA use this GDS?**

Guidance, by definition, represents FDA's current thinking on a topic; therefore FDA may revise a GDS as deemed appropriate before issuing it as an FDA guidance document. FDA defines guidance documents as those prepared for FDA staff, regulated industry, and the public that describe the agency's interpretation of or policy on a regulatory issue.

Therefore, the FDA presents this GDS to the advisory panel for input to assist in developing a framework for a future guidance. It is important to note that devices with a novel design, new material or indication for use may or may not be included in the guidance eventually developed from this meeting. If not specifically included, a novel design would be required to provide valid scientific evidence, such as a randomized clinical trial to demonstrate the safety and effectiveness or to support a specific device claim.

### **What is the Panel's charge?**

The panel will be asked to make a recommendation on the adequacy of the objective clinical endpoint criteria described for the proposed duration described in this GDS to support the safety and effectiveness or specific device claim(s) in a future marketing application. While the proposed GDS provides many elements necessary for a clinical study guidance document (patient/study success criteria, a proposed duration, etc.), we are concerned that other elements do not appear to be included. In order to move forward toward an appropriate guidance for clinical studies for hip replacement systems, we will ask the panel to consider what was proposed in the GDS, to discuss acceptable specific endpoint criteria and acceptable endpoints at a specific point in time, and to identify additional information that should be included for a future guidance on the topic of hip replacement systems.

### **Summary:**

This GDS proposes a standardized minimum patient success benchmark, based on 3 “variables” (Harris hip score (HHS), revisions and device-related complications), for prospective clinical trials that are intended to demonstrate the safety and effectiveness of hip replacement devices intended to treat hip joint disease that has resulted in pain and dysfunction in patients.

This benchmark was developed by clinical consensus by a group of orthopaedic surgeons who specialize in total hip joint replacement surgery. The consensus is based on their cumulative practice experience and a literature search aimed at discovering the failure modes of total hip joint replacement devices recorded in clinical studies of various hip replacement system (HRS) designs. The literature search was conducted using MeSH headings and sorted by the “Levels of Evidence,” a hierarchy of scientific evidence described by the Journal of Bone and Joint Surgery<sup>8</sup> that categorizes research studies in terms of validity and reliability of evidence.

### **Review:**

This document is intended to facilitate study design and provide consistency that will provide a least burdensome approach for review and introduction of new devices to market. It is focused on patient and study success, and only general areas of study design.

The GDS does not include a description of how studies were included or excluded and what the results of the meta-analysis were in terms of numbers of patients, length of follow-up, implant survival, revision rate, adverse event rates, and outcomes of assessment scales for pain, function and global satisfaction

FDA acknowledges the efforts of the team of expert orthopaedic surgeons who were charged with answering six questions to arrive at a consensus for the proposed benchmarks. Although the document addresses a list of complications and a minimal acceptable HHS, it appears to only partially address the minimally acceptable percent difference in clinical success as defined in the GDS between an HRS composite endpoint under study and the proposed consensus benchmark composite endpoint at the proposed time of the final postoperative evaluation. Also, the GDS does not appear to address what the minimal acceptable change in HHS at the final postoperative evaluation in comparison to baseline (question #2, GDS page 6) and the measures to be performed on postoperative radiographs with the minimal acceptable quantities of those measures (question # 4, GDS page 6).

### **Primary Endpoint Considerations and study duration**

Patient success is defined in terms of three primary endpoint outcomes: complications, revisions and HHS outcome at one year. Patient success is attained when a subject meets the quantities defined in all variables. Each patient can be deemed either a “success” or “failure.” The proposed standardized quantities for the variables required for patient success at one year are:

- A. Device related complications = 0 %
- B. HHS at 12 months greater or equal to 80
- C. Revision surgeries = 0 %

Study success is achieved when 95% of patients are deemed patient successes at one year.

While the absence of revisions and device-related complications is expected at one year, an HHS score of 80 may be too low a target for a one year time point. We are aware of large studies evaluating hip replacement devices report HHSs of 90 or better early in patient followup.<sup>4,5</sup> Thus, we will be requesting recommendations in respect to pain and function objective performance criteria (e.g., the appropriateness if the proposed HHS of  $\geq 80$  at one year or some other level for an objective performance criteria at another time point.).

Although the GDS does not specify which evaluations are repeated at which interval, the sponsor proposes data collection postoperatively at baseline, 6 weeks, 6 months, and 12 months. When queried in the past, the panel has previously recommended that 2 year data is not sufficient to understand the performance and safety of a joint prosthetic, even judging that 2 year data was an inadequate surrogate for long term performance. FDA has always recommended that data be collected at a minimum of 2 years and to collect data until the last enrolled patient has had his 24 month evaluation, as this allows for the collection of some data beyond 2 years. This GDS does not appear to have provided an adequate rationale or data to support the proposed objective performance criteria at 1 year. We will be asking for recommendations regarding the combination of the proposed objective performance criteria and the proposed time of the last evaluation.

### **Statistical Plan**

*For a full statistical analysis review, please see the statistician's review.*

The advantage of having a predetermined sample size is that sponsors would be better able to financially plan financially clinical trials, presumably reducing the cost of research and development. The proposed statistical plan includes a determination of sample size based on a delta of 4% with a confidence interval, between 91.7 and 97.5. This would allow a 5% failure rate at one year. This proposal appears to be more liberal than several large consensus and meta analyses that demonstrate that implant survival is greater than 95% at 10 years for cemented components and approximately 98% at 5 year follow up for uncemented components<sup>4,5</sup>. Based on the Power calculation in Appendix V, if the observed rate was 90% or 93% success, the null hypothesis would not be rejected, but the lower confidence interval could be 85% - 89% which may not be consistent with what is reported in practice or in the literature. We will be seeking recommendations on whether the proposes composite success rates at earlier time points (e.g., one year, 18 months, 2 years, etc.) are acceptable based on long-term information regarding success rates.<sup>4,5,6</sup> This GDS does not appear to address the treatment and analysis of patients requiring bilateral treatment or patients who are lost to follow-up. Other issues not specifically addressed in this GDS include what the primary analysis will be and how covariates will be considered and/or addressed.

### **Outcome Evaluation tools and Secondary Endpoint Considerations**

Although the document has suggested that radiographic and other secondary endpoint considerations should be conducted, no specific evaluations are proposed. In addition, OSMA proposes that some hip replacement designs may not conform well to standard, accepted techniques without citing examples or why this would be true. OSMA proposes that radiographic failures would be reported without discussing how this assessment characterizes the failure of a device, or which scales are appropriate for use in the approach proposed in the GDS.

### **Adverse Events**

Appendix IV lists the most common complications determined as a result of review of the articles provided in Appendix III. Not all complications are addressed for all materials and designs, particularly those with newer bearing surfaces and designs. This includes ceramic head or liner fractures. The table does not include the denominator for the number of hip procedures, the types of devices (bearing surfaces and materials, designs) that were included in the analysis for this evaluation describes or the type of study the data comes from. In addition, in contrast to that proposed in the GDS, FDA would include surgical technique (labeling for the device) as part of the device-related events.

### **Discussion**

We will be seeking a recommendation regarding the proposed clinical objective performance criteria as an alternative to other forms of valid scientific evidence such as a randomized clinical trial. We will also be accepting recommendation(s) regarding items discussed above that should be included to be complete and useful for evaluating the safety and effectiveness of these HRS devices for the purposes of supporting marketing applications.

This GDS proposed by OSMA brings several factors to the forefront in support of the proposed approach. This approach, which employs objective clinical performance criteria, would be a singular allowance in orthopaedic devices due to the long history of consistent device performance reported in peer reviewed literature and by professional experience. It is supplemented by some well-developed bench, preclinical

and clinical performance evaluations in standard use that characterize these devices well. Continuing professional society surveillance of the reliability of prosthetic devices with periodic international professional society meetings also exists which is dedicated to the evaluation of hip prostheses as the technology evolves. Moving forward with this equivalent of objective performance criteria can encourage the development and testing of new technology in the advance of the treatment of hip arthritis in a timely manner.

With a historical control, as proposed in this GDS, the resulting study becomes a one-armed observational study and since there is no randomization, any comparative statistical inference is compromised. An advantage of a randomized trial design is that confounding factors, such as selection biases, are counteracted because of the randomization. If there is no randomization, there is a greater need to check for and avoid potential confounding factors that lead to bias in the trial. Thus, there need to be tools in place to mitigate the biases that exist inherently with historical controls. The use of historical controls and a study of the past assume that the knowledge gathered can answer new clinical questions. The assumption is that a review of the literature allows complete and adequately detailed records for review. A meta-analysis based on published reports is subject to publication bias as negative clinical studies are less likely to be published. With a vast historical experience based on the patients and medical thinking of the past, there is a strong possibility of temporal bias. Baseline conditions of a population change over time. What naturally follows is a question of the capacity and soundness of a comparison between historical controls and future patients to be treated. The concern associated with those factors that have changed since the time of the historical review is whether the historical criteria applied to a new device may not discern whether the device is inferior to concurrent treatment.

All of the factors discussed above should be considered in the discussion of the GDS as proposed and presented to the FDA in order to effectively move a future guidance for the clinical evaluation of hip replacement systems forward.

### **References:**

1. Dawson, B, Trapp, RG, Basic and Clinical Biostatistics 3<sup>rd</sup> ed., New York: Lange, 2001, pp. 7-23, 63-72, 250-3.
2. Guyatt, G Rennie D, eds., Users Guides to Medical Literature: A manual for Evidence based Clinical Practice, Chicago: AMA Press, 2002, p. 426.
3. The Least Burdensome Provisions of the FDA Modernization Act of 1997: Concept and Principles; Final Guidance for FDA and Industry. 9/30/2002  
<http://www.fda.gov/cdrh/ode/guidance/1332.pdf>
4. Total Hip Replacement, NIH Consensus Statement, September 12-14, 1994, 12(5): 1-31.
5. Malchau H, Herberts P, Eisler T, Garellick G, Soderman P, The Swedish Total Hip Replacement Register., J Bone Joint Surg Am., 2002, 84-A Suppl 2:2-20.
6. Weinstein, J. The Dartmouth Atlas of Musculoskeletal Health Care, Chicago, IL: AHA Press, 2000.
7. National Center for Health Statistics, 1991 to 2000 National Hospital Discharge Survey.
8. Wright JG, Swiontkowski MF, Heckman JD, Introducing levels of evidence to the journal. J Bone Joint Surg Am, 2003, Jan; 85-A(1):1-3.