Designation: F3037 – 15

Standard Guide for Clinical Trial Design for Hip Replacement Systems (HRSs)[1]

This standard is issued under the fixed designation F3037; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide is intended as a resource for individuals and organizations involved in designing clinical trials of hip replacement systems (HRSs) including metal/polymer, metal/metal, metal/composite, metal/ceramic/polymer, metal/polymer/metal, and ceramic/ceramic bearing surfaces; semi-constrained and constrained designs; and cemented, nonporous uncemented, and porous-coated uncemented fixation.

1.2 In this guide, methods to measure the efficacy, effectiveness, and safety of HRS devices through standardizing outcomes measures are provided for designing, reviewing, and accepting human clinical trial protocols.

1.3 This guide is intended to provide consistency in study design, review, regulatory approval, and coverage approval for hip replacement systems to the health care market.

1.4 For the purpose of this guide, an HRS is any device that is intended to replace the hip joint, in part or in total, as a treatment for joint disease, trauma, or dysfunction, where long-term functional restoration and pain relief without major adverse events are the desired outcomes.

1.5 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

F561 Practice for Retrieval and Analysis of Medical Devices, and Associated Tissues and Fluids

F2809 Terminology Relating to Medical and Surgical Materials and Devices

F2978 Guide to Optimize Scan Sequences for Clinical Diagnostic Evaluation of Metal-on-Metal Hip Arthroplasty Devices using Magnetic Resonance Imaging

F2979 Guide for Characterization of Wear from the Articulating Surfaces in Retrieved Metal-on-Metal and other Hard-on-Hard Hip Prostheses

2.2 *ISO Standards*[3]

ISO 12891-1 Retrieval and analysis of surgical implants – Part1: Retrieval and handling

ISO 12891-2 Retrieval and analysis of surgical implants – Part 2: Analysis of retrieved surgical implants

ISO 14155 Clinical investigation of medical devices for human subjects – Good clinical practice

ISO 14971 Medical devices – Application of risk management to medical devices

3. Terminology

3.1 Unless provided in 3.2.1 – 3.2.5, definitions shall be in conformance with Terminology F2809.

3.2 *Definitions:*

3.2.1 *coverage, n*—insurance decision to reimburse for a device and/or procedure.

3.2.2 *effectiveness, n*—extent to which medical interventions achieve health improvements in real practice settings.

3.2.3 *efficacy, n*—extent to which medical interventions achieve health improvements under ideal circumstances.

3.2.4 *level of evidence*—strength of clinical evidence for evidence-based medicine (**1**)[4].

3.2.5 *safety*—the condition of being protected from or unlikely to cause risk or injury. See Appendix X1 for a tabulated list of adverse events reported for hip replacement systems (2).

3.3 *Acronyms:*

AJRR—American Joint Replacement Registry
ASA—American Society of Anesthesiologists
DVT—Deep Vein Thrombosis
EQ-5D—European Quality of Life – 5 Domains
FDA—Food and Drug Administration
HHS—Harris Hip Score

[1] This test method is under the jurisdiction of ASTM Committee F04 on Medical and Surgical Materials and Devices and is the direct responsibility of Subcommittee F04.39 on Human Clinical Trials.

Current edition approved June 1, 2015. Published August 2015. DOI: 10.1520/F3037-15.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

[4] The boldface numbers in parentheses refer to a list of references at the end of this standard.

HOOS—Hip dysfunction and Osteoarthritis Outcome Score
HRQL—Health-related quality of life
HRS—Hip Replacement System
ICD—International Classification of Diseases
LEAS—Lower Extremity Activity Scale
MCID—Minimal clinically important difference
MRI—Magnetic Resonance Imaging
OHS—Oxford Hip Score
PRO—Patient-reported outcome
PROMIS—Patient-Reported Outcomes Measurement Information System
QALY—Quality adjusted life year
RSA—Radiostereometric analysis
SAE—Serious adverse event
SF-36—Short Form (36 questions)
SF-12—Short Form (12 questions)
SF-6D—Short Form (6 dimensions)
THA—Total hip arthroplasty
TUG—Timed up and go
UCLA—University of California at Los Angeles
UTI—Urinary tract infection
WOMAC—Western Ontario McMaster Osteoarthritis Index

4. Summary of Guide

4.1 It is the intent of this guide to provide an overview of appropriate outcomes that are to be addressed in human clinical trials of hip replacement systems (HRSs). Depending on the requirements of the clinical trial, the outcomes to be addressed include hip-specific patient-reported outcomes, health-related quality-of-life patient-reported outcomes, activity level scales, gait speed, symptom relief (pain visual analog scales), and frequency of adverse events.

4.2 In general and in accordance with evidence-based medicine principles, patient-reported outcomes should be given preference over mixed outcome measures (surgeon and patient completion), intermediate outcomes (physical examination findings), or radiographic outcomes. However, the U.S. Department of Health and Human Services and/or local requirements may require mixed outcomes measures.

4.3 Because of the broad range of indications for HRSs, patient comorbidities, and functional/activity levels, it is impossible to identify or specify a single instrument score that measures the "success" of HRSs. Instead, a clinically significant improvement (minimum clinically important difference [MCID]) in a joint-specific, disease-specific, or quality-of-life instrument should be used as a measure of clinical "success" (30). A practical guide for determining MCIDs is that the MCID equals one half of the standard deviation of the change in the instrument score, MCID = $\sigma\Delta/2$ (**3**). This distribution method of determining MCID for a validated PRO instrument allows the calculation of the MCID for specific patient subgroups and/or interventions/treatments because the MCID may vary by patient subgroup and/or intervention/treatment.

4.4 The application of this guide does not guarantee clinical success of a finished product but will help to ensure consistency and adequacy in the clinical data of the clinical trial protocol.

4.5 The coverage criteria for medical treatments include: *(1)* that a net health outcome is achieved, *(2)* the clinical trial results are applicable (generalizable) to the patient population, and *(3)* the clinical trial results are applicable (generalizable) to medical providers (effectiveness versus efficacy). Therefore, clinical trials should be able to perform subgroup analyses based on patient characteristics (age, sex) and provider characteristics (community providers).

4.6 This guide does not suggest that all outcome instruments be used for each HRS. However, inclusion of an outcome measure from each section will provide a thorough description of the benefits of an HRS, including hip function/pain relief, health-related quality of life including a health utility measure with the ability to calculate Quality Adjusted Life Years (QALYs) (**4**), activity level, and mobility.

5. Significance and Use

5.1 Approximately 300,000 primary total hip arthroplasties (THAs) and 50,000 revision THAs are performed in the United States annually (**5, 6**). In addition, approximately 50 % of the 300,000 hip fractures in the United States annually are femoral neck fractures. The majority of femoral neck fractures are treated with hip hemiarthroplasties (femoral head replacement only).

6. Use (Outcome Measures)

6.1 *Patient-Reported Outcomes (PROs):*

6.1.1 Patient-reported outcomes (PROs) are vital to understanding the value patients receive from health care. Value can be defined as the change in quality of life and function divided by the total cost of care. Improvement in quality of life is most commonly measured by Quality Adjusted Life Years (QALYs) (**4**). QALYs are required for cost-effectiveness analyses and comparative effectiveness analyses used in coverage decisions. Standardization of PRO measures is necessary to compare outcomes of procedures (**7**). Standardizing PRO measures for implant and outcome registries will make comparative effectiveness data available to the clinical and regulatory communities.

6.1.2 *PRO Measure Selection*—PRO measure selection shall be pragmatic. High-respondent burden (too many questions) will result in poor rates of patient completion. High licensing fees make it difficult for not-for-profit registries to license the measure.

6.1.3 *Hip-Specific or Disease-Specific Outcome Instruments*—The hip-specific PRO measures most frequently used are the Oxford Hip Score (OHS) (**8**) and Hip dysfunction and Osteoarthritis Outcome Score (HOOS) (**9**). The OHS is used in the New Zealand Joint Registry (**10**) and the National Joint Registry of England, Wales, and Northern Ireland. The Oxford Hip Score and Oxford Knee Score have also been adopted for use in the United States. The American Joint Replacement Registry (AJRR) accepts the Oxford Hip Score and Oxford Knee Score as Level 3 data on patient-reported

outcomes. The Oxford Knee Score is the PRO knee instrument mandated by the Minnesota Department of Health for all knee arthroplasty procedures in the State of Minnesota effective January 1, 2012 (**11, 12**). The Western Ontario McMaster Osteoarthritis Index (WOMAC) (**13**) is a lower extremity osteoarthritis disease-specific outcome instrument used for lower extremity osteoarthritis. The Harris Hip Score (HHS) is a surgeon-reported outcome completed with patient input and subject to surgeon bias. However, because of the clinical and regulatory experience with the HHS, the HHS may be used as an outcome measure, but is not a preferred outcome measure due to potential surgeon bias. Pynsent et al (**14**) reviewed validated hip PRO instruments. A more recent validated PRO for osteoarthritis is the Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function instrument (Broderick JE, 2013).

6.1.4 *General Health-Related Quality of Life (HRQL) Outcome Instruments*—The European Quality of Life (EQ-5D) is used by the British National Health Service and National Joint Registry of England, Wales, and Northern Ireland to assess the HRQL change after THA (**15**). EQ-5D is used by the Swedish Hip Registry (**16, 17**), the Norwegian Arthroplasty Register (**18**), and the Norwegian Hip Fracture Register (**19**). The EQ-5D has become the standard outcome instrument for hip fracture outcome studies looking at femoral neck fractures (**20**), intertrochanteric fractures (**21**), and subtrochanteric fractures (**22**). EQ-5D is used for musculoskeletal disease research in Japan (**23**), Denmark (**24**), the Netherlands (**25**), and Finland (**26**). SF-36 and SF-12 are frequently used as HRQL outcomes instruments. However, the quality-of-life summary measure (SF-6D) is a calculated summary score and does not allow patient preference weighting for calculation of change in HRQL Both the EQ-5D and SF-6D can be used to calculate QALYs for cost-effectiveness or comparative-effectiveness analyses. The PROMIS Global Health instrument may be used to assess health-related quality of life (Amtmann D, 2011).

6.1.5 *Activity Level Scales*—The University of California at Los Angeles (UCLA) Activity Scale and Lower Extremity Activity Scale (LEAS) (**27**) were found to be the most valid activity scales for hip osteoarthritis (**28**).

6.1.6 *Gait Speed/Mobility Measures*—A significant percentage of patients with hip fractures have cognitive impairment and are unable to complete PRO instruments. However, gait speed can be measured independently of cognitive function if the subject is ambulatory. Also, gait speed (**29**) and mobility disability (**30**) are strong predictors of overall mortality. Therefore, measurement of gait speed change is a functional outcome measure for HRS. A standardized test to measure gait speed is the Timed Up and Go (TUG) test (**14**).

6.2 *Safety:*

6.2.1 Adverse event rates are a measure of safety and should be defined by the study protocol. All adverse events shall be recorded. Adverse events directly related to the HRS or otherwise required by regulatory guidance shall be reported. The following types of adverse events have been reported for HRS (**2**) and an example of how to report the data is included in Table X1.1. Additional adverse events that should be reported are: pseudotumor, adverse local tissue reaction, noise (grinding, clicking, popping, squeaking), taper wear, and increase in metal ion/corrosion products. Time windows for adverse event reporting should be based on regulatory guidance (Clinical Data Presentation for Orthopedic Device Applications, Food and Drug Administration, December 2, 2004). Adverse event reporting may be reported and analyzed according to both: *(1)* regulatory requirements and *(2)* time windows included in this guide in order to capture all adverse events and determine if different time windows affect adverse event rates.

6.2.2 Adverse event collection, analyses, and reporting protocols for a priori grading of adverse event severity and relatedness shall be established. An independent Data Safety Review Board should be considered when appropriate. A Clinical Events Committee should be considered when appropriate.

6.2.3 The following clinically expected events should be reported separately as hospital and/or surgeon quality measures:

6.2.3.1 Hip joint dislocation any time postoperatively;

6.2.3.2 Deep infection requiring re-operation within one year of surgery;

6.2.3.3 Deep vein thrombosis or pulmonary emboli or both within 90 days of surgery;

6.2.3.4 All-cause non-elective 30-day hospital readmission;

6.2.3.5 Intraoperative or postoperative femoral or acetabular fracture occurring within one year of surgery;

6.2.3.6 *Hip Reoperation/Revision Surgery (No Tiime Limit)*—A revision surgery is defined as a procedure that is performed on the replaced hip to remove and/or replace any femoral, acetabular, or both component(s) implanted at the index operation; or reduction of a dislocated hip replacement;

6.2.3.7 All Serious Adverse Events (SAEs), and

6.2.3.8 Death within 30 days of surgery.

6.2.3.9 For rare severe adverse events, consider increasing the level of significance (α).

6.3 *Radiographic Outcome:*

6.3.1 Radiographic analysis should be conducted. Measurements made on radiographs to determine implant position/migration are standardized in the literature (for example, Gruen zones and DeLee/Charnley zones). However, some HRS designs may not conform to these measurement techniques well. In such situations, alternative measurement techniques should be proposed by the sponsor. In either case, the sponsor should propose the definition of "radiographic failure" and report the number of failures.

6.3.2 Magnetic resonance imaging (MRI) should be used, when appropriate, to evaluate pseudotumors and soft tissues in accordance with Guide F2978.

6.4 *Wear and Other Radiographic Measures:* Radiostereometric analysis (RSA) and/or other radiographic methods may be used for measuring wear and implant stability/migration relative to bone.

6.5 *Retrieval Analysis*—Retrieval analyses should be conducted in compliance with Practice F561 and Guide F2979 and ISO 12891 Parts 1 and 2.