(i) Introduction and follow-up of new implants

Sanjeev Sharma
Kimberly Lammin
Peter Kay

Abstract

The recent recall of Articular Surface Resurfacing implants (ASR, DePuy Orthopaedics, Warsaw, Indiana) has yet again triggered a debate on the regulation of introduction and follow-up of new/modified orthopaedic implants in the market. Although the National Joint Registry was helpful in identifying the failing ASR implant, it took almost seven years from the introduction of the implant to the market to finally recalling it from the market. A comprehensive review of the systems needs to be done to prevent poorly designed implants from making it to the market and also to minimize the delay in identifying any potentially failing implants in use. The pre-market approval process (the CE [Conformité Européenne] marking in the United Kingdom [UK] and Europe) needs to be more stringent and be able to strike a balance between ensuring patient safety and promoting innovation. The post-market surveillance needs to be more effective to identify the failing implants early, and with the largest joint registry in the world, the UK is in an ideal position to provide this kind of surveillance. Better linkage with well-established regulatory authorities (The Orthopaedic Data Evaluation Panel and the Medicines and Healthcare products Regulatory Agency) and the use of Patient Reported Outcome Measures can make this process more effective.

Keywords implants; post-market surveillance; pre-market approval

Introduction

Joint replacement surgery is one of the greatest advances in the history of medicine. Ever since the development of total hip replacement by Sir John Charnley in the early 1960's,[1] millions of patients across the globe have benefited from joint replacement surgery, not only for hip disease, but also for disease involving other joints. All this has been possible due to a continual desire to innovate and improve. However, not all innovations have been success stories and we have witnessed some disasters related to new implant designs in the past few decades. In fact, even Charnley's initial attempt at total hip replacement using Teflon met with failure due to rapid wear.[2] Alternative bearing surfaces and implant designs have been developed in the recent past in order to reduce wear rates, extend survival and improve function. Sedrakyan et al conducted a systematic review of the literature recently to compare the effectiveness of various hip implant bearing surfaces and found no advantage for "hard on hard" bearings (metal-on-metal or ceramic on ceramic) compared with traditional "hard on soft" bearings (polyethylene based).[3] Also, in the recent past the use of larger size femoral heads has been promoted in primary hip arthroplasty to reduce the risk of dislocation[4] and enhance function.[5]

Failure of implants is a significant cause of disability, deformity and dysfunction to the patients, let alone the financial burden and increased workload that it imposes on healthcare systems worldwide. So far the focus has been on identifying the failing implants in order to stop their continued use in a timely fashion, and with this goal in mind the National Joint Registry (NJR) was set up in the UK in 2002. It does seem to have served its purpose to an extent, e.g. in identifying the failure of Articular Surface Resurfacing (ASR) implants (DePuy Orthopaedics, Warsaw, Indiana) including the ASR acetabular cup and the ASR XL femoral head, resulting in its recall from the market in August 2010.[6] However, before it was recalled in 2010, the ASR had been implanted in nearly 100 000 patients, and the result was a public health nightmare.[7] This raises two pertinent issues. Firstly, how to make the introduction of new implants more stringent without undermining innovation? And secondly, how to improve post-market surveillance to identify the potentially failing implant designs early on so they can be withdrawn from the market in a timely manner? In this article we review the current systems already in place to achieve these objectives along with possible ways to improve these further for the future.

Introduction of new implants

CE marking

Currently, for a medical device to be marketed in the European Union (EU), it must have a CE (Conformité Européenne) mark. CE literally means "European Conformity". The CE mark is an indication that a product complies with the essential requirements of applicable directives and that this has been proven in a conformity-assessment procedure. Moreso, this promotes a single European market where a CE marked product (medical or non-medical) can be freely marketed without restrictions. The essential requirements give particular consideration to:

- Technical performance – mechanical testing
- Safety of the device – compliance with international standards, and
- Medical performance – biological testing and data from clinical trials.

The conformity-assessment procedure is used to show proof that the requirements concerning safety and technical performance have been fulfilled, while medical performance is verified in the context of clinical assessment. The manufacturer has to obtain the CE mark from one of the several notified bodies

Sanjeev Sharma *MRCS FRCS (Tr & Orth) Arthroplasty Fellow, Wrightington Hospital, Hall Lane, Appley Bridge, Wigan, Lancashire, WN6 9EP, UK. Conflicts of interest: none.*

Kimberly Lammin *MRCS FRCS (Tr & Orth) Arthroplasty Fellow, Wrightington Hospital, Hall Lane, Appley Bridge, Wigan, Lancashire, WN6 9EP, UK. Conflicts of interest: none.*

Peter Kay *FRCS Consultant Orthopaedic Surgeon, Wrightington Hospital, Hall Lane, Appley Bridge, Wigan, Lancashire, WN6 9EP, UK. Conflicts of interest: none.*

(there are 83 such bodies in the EU[6]), and once the CE mark is obtained the device can be used throughout Europe without any hindrance. The devices requiring a CE mark are classified into Class I, Class II (a & b) and Class III, with higher classes signifying higher risk devices.[8] Class III is set aside for the most critical devices for which explicit authorization is required with regard to conformity for them to be placed on the market. Devices that are not classed as high risk do not require human clinical investigations prior to pre-market approval. In 2005 hip, knee and shoulder prosthesis moved from Class II b to Class III.[9] In the United States, the Food and Drug Administration (FDA) has a similar role in the pre-market approval, and it requires the manufacturers to submit their product to clinical testing to prove that it is both safe and effective for its intended use.[10] However, there is a fast track route as well, whereby a device (like the ASR) can be cleared by an FDA process known as 510(k), which refers to the section of the 1976 Medical Device Amendments to the Federal Food, Drug, and Cosmetic Act that created it. Under that section, the criterion for clearance of a new medical device is that if it is "substantially equivalent" to an already-marketed device (a "predicate"); clinical data are not required.[7] Such devices fall under the so-called "me-too group".[11] 90% of devices in the US are approved through the 510(k) route. Not surprisingly therefore, most of the problems in the US have been with devices approved through the 510(k) route. Between 2005 and 2009, 113 devices were deemed to be high risk and were recalled by the FDA, and 71% of these had been cleared through the 510(k) route.[12]

One of the common arguments that the manufacturers use to obtain a CE mark for a new implant is that it is similar to a device with a good track record. We know from past experience that even the so-called minor changes in implant design can significantly affect clinical performance. The Capital Hip (3M), which was marketed as a much cheaper version of the Charnley stem, was one the "me-too group" implants. The investigation led by the Royal College of Surgeons of England found several design features that potentially contributed to the failure.[13] Some of these stems were made of titanium rather than stainless steel, which allowed more bending and torsional micromovement of stems within the cement mantle. Shot blasting of stems produced a rougher surface in titanium stems compared to stainless steels stems. A combination of these factors resulted in excessive abrasive wear at the metal stem—cement interface. The proximal flanges were replaced with a triangular wedge of metal with rounded off edges, providing minimal resistance to torsional forces. The altered flange and the unwise use of the rasps (which were oversized by only 1 mm) also resulted in a thinner cement mantle proximally and distally, which in combination with a distal cement centralizer caused a defect in the proximal cement anteriorly. All these modifications led to a catastrophic early failure of these hips and their production was finally stopped in 1997. The matt finish Exeter stem,[14] Sulzer Inter-op acetabular shell[15] and Boneloc cement[16,17] are other examples of design failures that resulted in their withdrawal from the market. Therefore, CE marking of devices merely on grounds of equivalence with similar devices is inappropriate. In the absence of good data on clinical performance, the safety and risks of the device may be unclear at the time of introduction to the market.

Role of National Institute of Health and Clinical Excellence (NICE) guidance

The medical devices agency (MDA, now merged into the Medicines and Healthcare products Regulatory Agency [MHRA])[18] has issued guidance on the requirements for the introduction of implants used in hip joint replacements in the market. This is largely based on the guidance issued by NICE in the UK. They suggested that although the NICE guidance applied solely to primary total hip replacements, manufacturers could consider similar criteria for other joint replacement implants as well.

The guidance defines the "benchmark" effectiveness of primary total hip replacement implants as "a revision rate of 10% or less at 10 years". Furthermore, the guidance states that "evidence used in support of any prosthesis, to establish whether or not it achieves this benchmark, should relate to data on 10 or more years follow-up from a number of centres, obtained via adequately sized, well conducted observational studies (preferably with consecutive patients from non-selected populations) or randomized controlled trials". Given that there are several different designs on the market, it is impractical to have randomized controlled trials for all prostheses.

The NICE guidance also defines an "entry benchmark". This applies to an implant that does not meet the 10-year benchmark, but that does have 3-year data collected as above, demonstrating "performance consistent with the benchmark of a 10% revision rate at 10 years". The NICE guidance states that such prostheses need to be subject to annual review (up to 10 years) "to ensure that the revision rate remains consistent with the 10 year benchmark".

Therefore, the minimum survival that is expected of a primary total hip replacement implant has been established by NICE. Survivorship should typically be 97% at 3 years, with a subsequent revision rate typically not exceeding 1% year-on-year thereafter (i.e. survivorship of at least 95% at 5 years and 90% at 10 years).

Manufacturers should construct any investigation or study protocol, whether in the pre-market or post-market phase, in such a way that the survivorship data can be established and compared with the benchmarks defined by NICE.

Orthopaedic Data Evaluation Panel (ODEP)

The Orthopaedic Data Evaluation Panel creates a database of all the products submitted to the panel by the industry in relation to all the NICE benchmarks.[19] For 10-year benchmark (Full Benchmark) products, ODEP place products in one of four categories:

- Level A − strong evidence, which signifies a failure rate of $\leq$10% in a cohort of >500 joints at the start of study. The data submitted could be from joint registry or a multicenter trial ($\geq$3 centres including a non-developing centre) and should include Kaplan—Meier survivorship at 10 years.
- Level B − reasonable evidence, which signifies a failure rate of $\leq$10% based on multicenter data (>1 centre including a non-developing centre).
- Level C − weak evidence, which signifies a failure rate of $\leq$10% in the studies submitted, but with poor quality data. They are given 2 years to improve their data, failing which they are deemed unacceptable.
- Unacceptable evidence.