Position paper from EFORT and NORE (The Network of Orthopaedic Registries of Europe) on:

"Quality assessment of Orthopaedic Implants"

Rob GHH Nelissen, Per Kjaersgaard-Andersen, Keith Tucker

February 19 2019

Quality assessment, its importance

A "Quality" assessment of registries and benchmarking can only be made by a group, independent of individual registries regulators, manufacturers and benchmarking organisations. EFORT/NORE is an organisation which is completely independent and therefore best placed to validate and oversee the European Arthroplasty registries and the benchmarking consortia in Europe.

This important function has become even more relevant with the introduction of the new Medical Device Regulation (MDR) which states that manufacturers will need contemporary clinical evidence to support maintenance of their CE marks for legacy products and for the introduction of new devices.

This evidence that the manufacturers require is available through many of the national/ regional Registries. The overall goal is patient safety.

Implants of interest

Initially the emphasis would be on Total Hip and Knee replacements. It is thought that the sum total of these operations, annually across Europe, is thought to amount to more than 1 million operations

Stakeholders

There are many stakeholders in "Quality Assessments". Obviously, the patients come first as this is all about making sure they get implants which will give them excellent function for a long period, preferably for the rest of their lives.

Other important stakeholders include:

- The Orthopaedic surgeons (who want to be sure the implant they are using is well tested and long lasting).
- Manufacturers who need to know that their implants are performing satisfactorily and besides others,
- Procurement agencies. In the UK, the NHS hospitals rarely procure non ODEP (Orthopaedic Data Evaluation Panel, vide infra) benchmarked implants. Many Healthcare insurance companies (in The Netherlands the direction of travel is towards at least NOV1B (ODEP 5A), NOV1A (ODEP 10A) benchmarked being available for general use)
- Hospitals.
- Regulators including the European Competent Authorities
- EFORT/NORE with the specialty European societies (EHS, ESSKA, SECEC)
- ODEP and NOV (Netherlands Orthopaedic Association, as member of ODEP)

International Benchmarking

Worldwide, there are two well established organisations who benchmark joint replacements. They are ODEP (The Orthopaedic Evaluation Panel established in 2002, http://www.odep.org.uk/) and NOV (Netherlands Orthopaedic Association)

1. ODEP benchmarks Total Hip, Knee and Shoulder implants at 3,5,7,10,13 years and benchmarks the data that is used in the manufacturer's submission with an A* (the best), A, B and unacceptable grading. ODEP is used globally. Since its inception ODEP has continued to evolve with the latest changes in the methodology being introduced in 2017 and with the 13 year benchmarks introduced in 2018.
2. NOV (Netherlands Orthopaedic Association) benchmark use the classification 1A (5 yrs), 1B (10 yrs), 2 (no evidence) and since 2016 members of NOV are also members of ODEP and attend all the ODEP meetings. In 2018, NOV adopted the ODEP methodology so as to achieve global uniformity, with the exception that if data from the Dutch Arthroplasty Register (www.LROI.nl) are worse than ODEP data for a particular implant the benchmark is downgraded to the LROI data in the Netherlands.

Other Organisations that have shown an interest in benchmarking

ISAR (International Society of Arthroplasty Registries) has drafted some suggestions for an "ODEP+" classification. This was discussed at the ISAR meeting in Iceland (2018). Compared with ODEP this system proposes:

- More rigorous statistical analysis of submitted data
- Larger cohorts, especially for the early benchmarks (e.g. 2 years)
- Submitted data being confined to observational registry data to the exclusion of data from presentations, publications, in house data etc
- A 2 year benchmark instead of one at 3 years,
- Dropping the 7 year benchmark
- No use of a 13 year benchmark

It should be remembered that, for the most part, early benchmarks can only pick up catastrophic failure. Furthermore, mean survival analysis data at early follow-up are likely to have very wide confidence intervals, stressing the uncertainty of these data.

The downside of demanding large cohorts of patients to be used for early benchmarks is that a manufacturers with limited access to the market, will take a very longtime to accumulate enough patients to submit data for a benchmark. This raises the possibility of ignoring mediocre results until it is far too late (i.e. a too large cohort has been exposed).

Implementation strategy

Education and spreading the word

If member states are going to use benchmarking to help their surgeons in the selection of the joint replacements they use it will only happen if they fully understand it (i.e. simplicity is more important than complexity).

They are more likely to embrace such a system if they have a degree of ownership for it. This will include using data from their own national registry (where this is available) as part of any assessment of a product.

Educational opportunities for surgeons by EFORT / NORE

- Making everyone aware of the advantages of benchmarking can be achieved through articles in newsletters and journals
- Presentations at EFORT and the European orthopaedic speciality society meetings
- Demonstrations of the process. ODEP for knees will be conducted during the EFORT congress in Lisbon 2019
- EFORT has the opportunity of "badging" a benchmarking system such as ODEP perhaps by using the logos ODEP^{EFORT} or ODEP^{EU}
- **ODEP^{EFORT}** classification could be prominent on the NORE part of EFORT website and on the front page of the EFORT website
- Employing an ODEP fellow (i.e. writing article, visiting registries) through EFORT fellowship programme.

Education by and for Manufacturers

Without the support of MedTech Europe, national manufacturers organisations such as ABHI in the UK, and the national orthopaedic associations representing orthopaedic surgeons, any expansion of benchmarking for implants is likely to fail.

Manufacturers are used to ODEP (and NOV in the Netherlands) and it is widely known that they are all in favour of one global system like ODEP where they see the classifying metrics are succinct and objective. Therefore, the Netherlands has in 2018 adopted ODEP for there implant classification, with the exception on national implant performance (vide supra).

Having representatives of EFORT, specialty societies, the implant manufacturers and European patient representatives on a committee to see the implementation of benchmarking across Europe would be essential.

Validation

Continuous validating a benchmarking system is both important and challenging if it is to be worthwhile.

- Where it is possible, referencing data from more than one source is clearly important
- Similar data from more than one registry will underscore the validity of a submission, given the heterogeneity that exist between countries (i.e. indication, surgical technique, phenotype of patients etc)
- Validation is also achieved by a manufacturer being obliged to resubmit for a higher benchmark every 2-3 years. If no data are submitted, the earlier benchmark classification is rejected.

Representation within benchmark committee

If Europe is to embrace benchmarking then an over arching committee should be set up to oversee its introduction and development

- ODEP committee should have two EFORT seats (1 NORE, 1 EFORT board) and a rotating seat of Specialty Society ((e.g. EHS, ESSKA SECEC etc) depending on the implant to be classified (e.g. Exeter **ODEP10A**$^{\text{EFORT}}$ a EHS seat etc).
- For each joint (Hips, Knees and shoulders) there will be a need to have a separate group to oversee consistency and the development of the initiative.

Funding

At present ODEP ratings are free to all manufacturers although the manufacturer will obviously incur expense putting together their submissions. Data for submission to ODEP can be requested by manufacturers from a number of existing registries. Two models of data delivery to manufacturers exits: delivery of raw implant data of that specific manufacturer (e.g. NJR model "Supplier Feedback") or aggregated data which are analysed (e.g. the Dutch LROI or the Australian model). Most registries charge an annual and/or per implant request fee.

With any expansion of the ODEP benchmarking model some funding would be needed, principally to fund travel and accommodation expenses for the panel members besides salaries for support staff.

Up until recently ODEP surgeon's expenses and secretarial expenses have been paid from the British taxpayer via the NHS supply chain run by DHL. This has recently changed and moving forward SCCL (Supply Chain Co-ordinator Limited) will be responsible for expenses. The expanses for the Netherlands representatives to ODEP are reimbursed by NOV. In future the funding model may need to change, possible options are:

- Apply for funding through Horizon2020 at EU for travel/hotel/lunch costs ODEP panel meetings.
- Reach out to European patient alliance
- Reach out to EU commission
- EFORT HQ support staff

Potential criticism on ODEP guidelines (see earlier)

The ISAR benchmarking group (which is basically ODEP+) has criticised ODEP, with focus on methodology aspects.For that matter ODEP has reviewed and implemented basic benchmarking methodology in 2016.

Ideally, the highest ODEP classifications ODEP 10A* and beyond are based on more robust data sources, such as data from 2 to 3 validated regional or national registries. Such a new classification (i.e. **ODEP10A^{EFORT}**) could be used parallel to the existing classification. The latter is important since some smaller manufacturers (with potential good products) only supply a limited market, perhaps only in their own country. ODEP also accepts results from RCTs, peer reviewed publications and presentations. They also sometimes accept, usually as part of a submission "in house validated studies" for a rating, and these data are likely to be less robust. ODEP subscribes to the theory that early registry data are often unreliable and wherever possible best supported by clinical trials and RCTs

Where manufacturers use data from registries these registries must meet the ISAR standard for registries to become a full ISAR member (i.e. at least 85% completeness for primary surgery).

Implementation of implant benchmarking within EFORT community

Rolling out benchmarking in the form of ODEP, to all member states all at once would be problematic. There are untold numbers of problems that might be encountered. It is suggested that for implementing an EFORT wide system ODEP^{EFORT} the focus should be on one joint to start with and the hip, would be most feasible. It would start in two or three countries. For that matter ODEP benchmarking is used in the UK/Wales and The Netherlands for both hips and knees. Addition of a third country or European region in the starting phase would create momentum. After one year the ODEP^{EFORT} for knees could start in these three countries and a fourth hip country or European region can start etc. Since ODEP is implemented for hips, knees and shoulders in two countries, one could argue to start with hips and knees ODEP^{EFORT} in three countries.

Another option could be to have an EFORT^{HIP} and EFORT^{KNEE} etc classification, identifying those implants which have at least 95% survival at 10 years. Both ODEP and registry data are input for such an EFORT classification, but (validated) registry data are the leading source for such a classification.

Introduction of new implants

Performance of new implants depends on implant-bone fixation and surgical technique.

Poor implant fixation leads to implant loosening early (five years) or late (10 years). However, there is now 40 years analysis of implant micromotion techniques (RSA: radiostereometry) available which has a predictive value within 1-2 years for both early or late (10 years) loosening. It is strongly argued that as the RSA technique is so highly accurate, only 50 patients need to exposed to a new innovative implant design to confirm (or not) that fixation has been achieved with a new design. The qualification **ODEP^{RSA}** (i.e. RSA data for 2 years) would underscore the importance of patient safety and still stimulate implant innovation by preventing implant disasters to patients.

The second impact on implant failure is surgical technique. In that respect, orthopaedic surgeons are the only stakeholders that evaluate and analyse, the effects of surgical technique (i.e. surgical exposure, instrumentation for implanting a prosthesis) on implant performance (i.e. implant migration and loosening). Therefore, in the parts of the UK served by the NJR , "Beyond compliance" has been introduced, to try and reduce both the implant and surgeon risks associated with the introduction of a new implant. "Beyond Compliance" is designed to improve the safety of the introduction of new implants by undertaking a risk analysis, closely monitoring and reviewing its performance with extra data (like adverse events, X Rays etc) being added to the standard implant data in the NJR database. Two of BC's functions include user group meetings and direct mailing of all surgeons who have used these products, (and again, if they have been revised). It is intended to hold "User group meetings" for surgeons who have been using BC products inside and outside the UK at the EFORT meeting in Lisbon