

# Group A: Using a registry – long-term follow-up provides relevant evidence

## Judge A Handout

**Using a registry: How long term follow-up provides relevant evidence**

Andrew Judge

Professor of Translational Statistics  
Director MSc Orthopaedic Surgery  
Musculoskeletal Research Unit, Translational Health Sciences,  
University of Bristol, Bristol, UK

**Introduction**

- How did the work done in relation to joint prostheses identify appropriate evidence that has directly influenced policy and practice?
- What possibilities are there for monitoring outcomes and adverse events in implants in general?
- How can medical implants be evaluated in an evidence-based way?
- Is there a science/statistical approach to monitoring registry reports in real time to identify anomalous developments?

**Scott's parabola**

"Scott's parabola" (produced from Scott, JW. Scott's parabola: the rise and fall of a surgical technique. *BMJ* 2011;323:1477)

**The National Joint Registry – NJR**

[To] collect relevant data about joint replacement surgery in order to provide an early warning of issues relating to patient safety!

- Started in 2003, to:
  - Improve quality of outcomes
  - Ensure the quality and cost effectiveness of joint replacement surgery
- 5 main surgery types
- >2 million joint replacements included

101,651 hip replacements  
108,713 knee replacements

2017

Fig 1. (A) illustrates an UK resurfacing implant, (B) ceramic on ceramic, large head, cemented acetabular components (cemented), and (C) ceramic on polyethylene, small head cemented acetabular components (with stem).

Range of stems, cups and combinations

Choice of implant combinations in total hip replacement systematic review and network meta-analysis

2008 – 2017

- 189 different stems – 72 first used in 2008
- 189 different cup/shells – 89 first used in 2008
- 2,119 unique combinations – 154 entries of new combinations

**NIHR** Digital Biomedical Research Centre

**Uptake of top 5 new cups**

**Uptake of top 5 new stems**

This study was supported by the NIHR Biomedical Research Centre at University Hospitals Bristol/NIHR Foundation Trust and the University of Bristol. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care

**Novel methods to improve the detection of outlying surgeons and units**

My performance is highlighted by the blue and black triangles.

Sources: Maggioni et al., *JAMA* 2011  
Hendon et al., *Stat Med* 2007

**The Bone & Joint Journal**

Temporal trends and geographical variation in the use of subacromial decompression and rotator cuff repair of the shoulder in England

Abstract

Background: The use of subacromial decompression (SAD) in the management of rotator cuff disease has increased significantly in the last decade. This study aims to describe the temporal trends and geographical variation in the use of SAD and rotator cuff repair (RCT) in England and Wales.

Methods: The number of SAD and RCT procedures performed in England and Wales from 2004 to 2014 was determined using data from the National Joint Registry (NJR). The number of SAD and RCT procedures performed in each region was also determined.

Results: The number of SAD procedures performed in England and Wales increased from 1,000 in 2004 to 10,000 in 2014. The number of RCT procedures performed in England and Wales increased from 5,000 in 2004 to 15,000 in 2014.

Number of SAD procedures  
Number of RCT procedures  
Total number of SAD and RCT procedures

Geographical variation in the use of SAD and RCT procedures in England and Wales.

**THE LANCET**

Arthroscopic subacromial decompression for subacromial shoulder pain (SAD): a multicentre, pragmatic, parallel-group, placebo-controlled, three-group, randomised surgical trial

Background: The use of SAD in the management of rotator cuff disease has increased significantly in the last decade. This study aims to describe the temporal trends and geographical variation in the use of SAD and rotator cuff repair (RCT) in England and Wales.

Methods: The number of SAD and RCT procedures performed in England and Wales from 2004 to 2014 was determined using data from the National Joint Registry (NJR). The number of SAD and RCT procedures performed in each region was also determined.

**Adverse outcomes after total and uncemented knee replacement in 105 839 matched patients: a study of data from the National Joint Registry for England and Wales**

Failure rates of metal on metal hip replacements: analysis of data from the National Joint Registry for England and Wales

Failure rates of cemented metal on metal hip replacements: analysis of data from the National Joint Registry for England and Wales

**Is the rate of revision of 36 mm metal-on-metal total hip arthroplasties with Pinnacle acetabular components related to the year of the initial operation?**

An interrupted time series analysis using data from the National Joint Registry for England and Wales

Primary observed rate  
Trend before 2007  
Trend 2007 onwards  
Expected observed rate  
Trend before 2007  
Trend 2007 onwards

**Adverse reactions to metal debris occur with all types of hip replacement not just metal-on-metal hips: a retrospective observational study of 2340 revisions for adverse reactions to metal debris from the National Joint Registry for England, Wales, Northern Ireland and the Isle of Man**

Number of revisions for adverse reactions to metal debris  
Metal-on-metal  
Metal-on-polyethylene  
Polyethylene-on-polyethylene

- Moderator: Leeza Osipenko, LSE
- Reporter: M Payne, HealthWatch
- Lynne & John Sharman of Sling the Mesh
- Andrea ... KCL Medical School
- Rachel Cooper of Transparency International Health Initiative
- Robert Grant, medical student, Leicester University

## First session

Presentation by Andrew Judge, Professor of Translational Statistics, University of Bristol:

The experience of what happens when a new device or implant comes onto the market was shown in a light-hearted paper in 2001. In “Scott’s parabola”, a device enters the market, usage escalates until it becomes the standard treatment and levels off, then the first reports of adverse effects come in, usage starts to fall, until widespread publicity and the device falls into disuse.

A possible solution for post-marketing assessment has been in place since 2003. HQIP (Health Quality Improvement Partnership - a publicly funded body for England, Wales, NI and Isle of Man) runs 40 registries, including the National Joint Registry. The NJR covers hip and knee replacement surgeries, of which there are currently records of over 100,000 of each in the registry to date. It captures data on each device, matched to patient and surgeon. The outcome data measured is simple: numbers of additional surgery needed to revise the initial operation. The system is easy to use and NJR now has virtually 100% compliance. There are approximately 50 registries in different countries of Europe.

## Discussion 1

### *Benefits:*

- NJR’s automated data generation can flag outliers – e.g. a surgeon or a unit with an unusually high fail rate. Surgeons and units receive feedback on their performance, motivates to maintain standards.
- Can use database to design more in depth research, designing methodologies with additional outcomes, e.g. pain, mobility levels. For example, a 10-year study of rotator cuff repairs (“CSAW”) could find no evidence that the surgery offered any benefit to the patients.
- Although the data is not fully publicly available, it is possible to apply for access to the data (unlike for pre-approval data, which is commercial in confidence and completely hidden).
- Countries too small to have their own registries can apply to access other country’s data, so to support public health decision-making on devices in their own countries.

### *Issues:*

- Private healthcare can show a lower fail rate, reflecting a bias towards going private for the most straightforward operations, while complex ones are dealt with by NHS.
- There are 5 different types of surgery covered, yet e.g. in hip surgery there are over 2000 unique combinations of the various stems and cups in the joints. Why so much innovation needed, if most patients already have good outcomes? New devices are adopted very quickly – some surgeons are particularly keen to try new things. There can be pressure from patients to “do something”, and surgeons like to operate. Patients like to think they are getting the latest technology.
- Knowing that these operations can have a life of 25 years, we don’t have retrospective information for people who had these implants before data was collected.
- May be resistance from other devices? In the case of hip and knee surgery, the operations have a high success rate, and the outcome is easy to measure. But devices with poorer outcomes would not look as good; and outcomes may be less clear or harder to define and measure, making it difficult to get consistent data measurement across different registries.
- It is still taking too long to pick up problems with devices. The data analysis finds patterns of poor outcomes retrospectively, but signals should be found earlier to enable action.
- Combining registries internationally would create a bigger data pool, even a global registry of all implants, but how to make the data consistent? How to incentivize data reporting? E.g. make reporting a mandatory part of surgical procedure.

- Funding and accountability – who should fill the registries information, who keeps the data current, how to define the outcomes to measure? Here, COMET (Core Outcome Measurement initiative exists for treatments, is only applied to some devices).

## B: Evidence and non-evidence synthesis

- Who regulates the registry? In the case of NJR it is HQIP and the NHS (Public Health England). But would they intervene effectively if compliance went down? Example: there used to be a registry for breast implants, but it was discontinued. When the PIP scandal came out, and they wanted to track surgeries for recalls, there were no records.
- Problem: if data is not collected then you can't monitor the problems. Yet each implant has an associated operation, so how difficult can it be to record what is put into each patient?
- If all devices are now to be identifiable, tracking the person who received faulty ones should be easy. But there is a huge IT challenge to be able to apply that information in context of a shaky NHS computer system.
- There can be too much data, it can be full of holes and therefore of not much use. It needs to be standardized, with standardized methods of evaluation, and many problems would then be picked up more quickly.
- Evidence generation is left to industry. No-one is responsible for delivering clinically. NICE cannot ask the questions, it can only process the available data and make recommendations based on that.

## C: Risk vs numbers – where do we concentrate?

- Biased information – it is left to the manufacturer to produce communications for the patient, and even for the surgeon. Even the surgeon has no independent data because there is none.
- Registries have the data, surgeons are under scrutiny, with peer pressure to perform better.
- Information on risk is not produced in the quantity and quality needed – Health Technology Authorities are not helping to inform decision-making. The Government does not stand up for patients' need to get unbiased information from doctors.
- Risk is related to use – in the case of a device, it is implanted, so you can't easily stop using it.
- The MHRA has 600,000 devices to monitor, not including apps and software.
- The MDR are written, but there are opportunities for HW to influence the way that the rules are implemented.
- Long term effects, and where effects are rare, you can mitigate against the problems but don't stop them.

## D: How can the healthcare community better support regulation for patient benefit?

- Who has responsibility/ownership of the problem? If a problem is found with one device, what can be done regarding similar devices to it? Why can licences not be withdrawn?
- Registries are relatively cheap to run, are getting cheaper, but still more expensive than simply reporting SAEs to regulators.
- We have to boil these down to simple messages! The MHRA is not a panacea, there are many moving parts in the whole system. It is far too complicated for patients.
- Need to move away from patients needing to sue after problems, to a no-fault compensation to solve patient problems much more quickly.
- Registries focus on long term outcomes. How can we collect data on patient-relevant outcomes more early?
- Manufacturers complain their competitiveness would be disadvantaged, but it would not if all of them were registered subject to the same regulation!
- We need to know better *who* we should be talking to, and need a few simple messages, to put to IDEAL, Cochrane, manufacturers?

# Group B: Evidence and evidence synthesis for non-randomised studies of medical devices and implants

## Reeves B Handout

### Evidence and evidence synthesis for non-randomized studies of medical devices and implants

1. Given the nature of implants and devices, what sources of evidence are likely to be available?
2. Are there any guidelines for judging the quality of available evidence? (or if not, should someone be developing them?)
3. How can studies be combined and compared?
4. What further work needs to be done to provide a good evidence base?
5. What can we do to promote appropriate developments in this area?

1

### Given the nature of implants and devices, what sources of evidence are likely to be available?

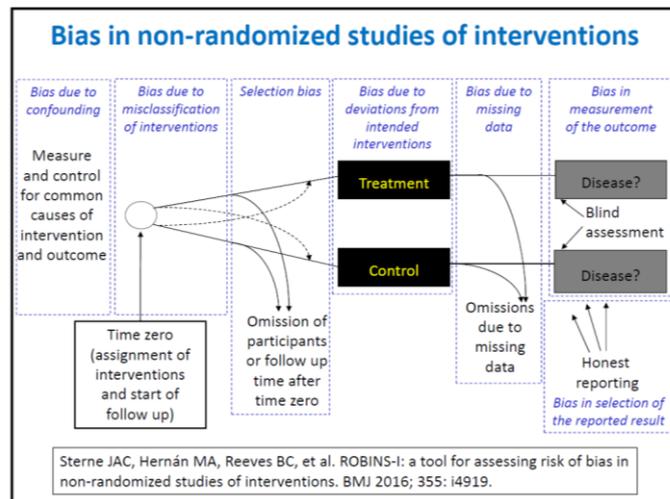
- Surgical setting (often)
- Existing devices work well (often), hence aim to detect a small incremental benefit, requiring a large sample size
- Outcomes of interest are often rare (but potentially large effects) or long term, yet device 'life cycle' may be short.
- Are process/surrogate outcomes (showing a device does what it claims to do) sufficient, or are 'clinical' outcomes (showing a device confers "good health"), required? (Analogy with RCTs of diagnostic technologies.)
- Evidence rarely comes from RCTs. Many sources of evidence are case series / registers.
- May or may not include 'control' (comparative vs non-comparative studies); beware eligibility criteria for registers.
- Studies to demonstrate equivalence are difficult to do ...

2

### Are there any guidelines for judging the quality of available evidence?

- ROBINS-I for comparative studies – seven dimensions of bias (next slide); each dimension judged to be at low risk, moderate risk, serious risk or critical risk (ignore).
- No established guidelines for judging quality of non-comparative studies. *What evidence do such studies contribute?*
- Beware conflicts of interest among evidence producers (including non-commercial studies)

3



4

### RCT (ISAT) versus register study (National Study of Subarachnoid Haemorrhage): clipping vs coiling

In intention-to-treat (ITT) analyses of RCTs, cross-overs shift the estimated treatment effect towards no difference.

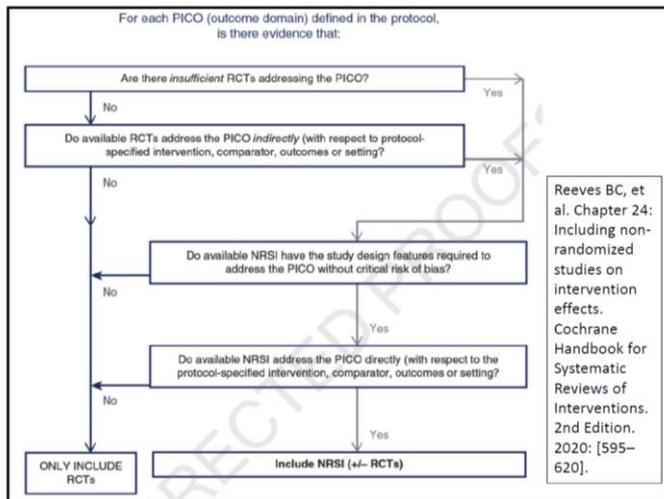
- In ISAT, 9 patients allocated to coiling actually had clipping and 39 patients allocated to clipping had coiling.
- In NSSH, information about intended treatment and cross-overs was not available and ITT analysis was not possible, because the intended mode of treatment was not documented.
- In ISAT, all randomized patients were assigned to coiling or clipping treatment and their data were analysed, even if they had neither. A procedure was not completed or not attempted in 4.9% of 2107 patients. Survival was known and analysed for all patients; the outcome of 'dead or dependent' was missing at 1 year for 1.1%.
- In NSSH, 9.3% of 2397 patients were not coiled or clipped, and were excluded. The combined outcome of dead or dependent was missing for 9.4% of patients who had coiling or clipping.

5

### How can studies be combined and compared?

- Apply usual systematic methods for reviews of the effectiveness of an intervention – but extract study design features, apply ROBINS-I and do not consider evidence at critical risk of bias. Difficult to do / time-consuming.
- Some evidence (collated through a systematic review) may be worse than no evidence (depending on the review authors / decision-makers ). **Discuss!**
- Do not consider case studies/series (non-comparative studies) evidence of effectiveness. These are at risk of other biases (not well researched). Comprehensive registers can provide descriptive estimates of outcomes, but not effectiveness. **Discuss!**

6



7

### Study features, not design labels

Table 2. Checklist question/item

	RCT	Q-RCT	IV	RD	CITS	ITS	DD	CBA	NRCT	PCS	RES	HCT	NCC	CC	XS	BA
1. Was the intervention/comparator: (answer "yes" to more than 1 item, if applicable)																
Allocated to (provided for) administered to/chosen by individuals?	P	P	Y	Y	P	P	P	P	P	P	P	P	Y	Y	P	P
Allocated to (provided for) administered to/chosen by clusters of individuals?	P	P	N	N	P	P	P	P	P	P	P	P	N	N	P	P
Clustered in the way it was provided (by practitioner or organizational unit)?	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P	P
2. Were outcome data available: (answer "yes" to only 1 item)																
After intervention/comparator only (same individuals)?	P	P	P	P	N	N	N	N	P	P	P	P	Y	Y	Y	N
After intervention/comparator only (not all same individuals)?	N	N	N	N	P	P	N	P	P	P	P	P	N	N	N	P
Before (once) AND after intervention/comparator (same individuals)?	P	P	P	P	N	N	N	P	P	P	P	P	N	N	P	Y
Before (once) AND after intervention/comparator (not all same individuals)?	N	N	N	N	P	P	P	P	P	P	P	P	N	N	N	P

Reeves BC, Wells GA, Waddington H. Quasi-experimental study designs series-paper 5: a checklist for classifying studies evaluating the effects on health interventions – a taxonomy without labels. Journal of Clinical Epidemiology 2017; 89: 30–42.

8

### What further work needs to be done to provide a good evidence base?

Suggestions (discuss):

- Registers – but existing national registers are imperfect (completeness/accuracy, e.g. National Joint Registry)
- Disease/condition or intervention register?
- Mandatory registers; mandatory recording by hospitals (doctors and coders) of device-related complications in HES; new OPCS4 procedure codes?
- Use bar code systems in theatre for easy data capture, linked with national database of device identities / serial numbers. (cf. Unique Device Identifiers, EUDAMED, Bruckner, p.8)
- Consider pros and cons of tighter regulation – fewer complications (greater vigilance) but longer time-to-market and higher price (cf. drugs), potentially (far) fewer devices (good/bad)? May depend on specific device / health condition being targeted).

9

### What can we do to promote appropriate developments in this area?

Suggestions (discuss):

- Don't give a device a CE-mark without company providing model / serial numbers of all devices sold in the country
- Require registration of studies (so evidence cannot be buried)?
- Advocate for better regulation. But what regulation?
- Be careful what you wish for ...

10

## Reporters Notes:

A key thread throughout the discussion was, "How do we emulate the trial we would like to do." It was clear that RCTs are rarely likely to be appropriate or ethical when considering what evidence for a medical device is adequate to sanction its use. In terms of studies that are more generally used, two categories were considered: comparative and non-comparative.

Problems with non-comparative studies were that they tell us nothing about whether they describe the best treatment or even one that is better than alternatives: comparisons with other studies are not possible. Comparative studies are required but not always carried out.

The difficulty in assessing the quality of non-comparative studies was highlighted.

Registries could be a good idea but not without problems. There are some that are already commonly used and work well — we need to learn from these. Need to ensure they collect appropriate and relevant data and are not burdened with irrelevant data that could obscure and confound vital analysis. It was noted that some registries have failed because they had been too ambitious.

Ideally, failure rates would be low and that means that can be difficult to detect so good systems of flagging up possible trends early on are required but the danger of false positives needs to be considered.

The costs of setting these up could be significant — it's not clear whether any existing registries could be used as a model and scaled: there are something like 600,000 different types of medical devices being used but the number of individual ones is not known. This, of course, requires that registers be updated systematically and conscientiously so operational hours are accumulated and endpoints with both successes and failures noted. The costs and complexities of setting this up, populating it, adding new data and maintaining it throughout device and patient lifetime should not be underestimated. But the wider issue is that there is no point in gathering data if it is never

actually used: there needs to be systematic ways of continually analysing these data to spot issues, ideally at an early stage. This should also consider careful analysis of possible trends to highlight issues (such as device failure after a number of operational hours) before they became widespread and systems in place to assess the hazard and risk. This will come at a cost.

A concomitant issue is individual device traceability. A representative of a device manufacturer said that all devices would be tracked and traced through a modern manufacturing process and subject to strict quality and statistical process controls. It was agreed that there should be a method by which all implantable devices could be associated with the patient in whom it was used. Some technology would be required but it would appear to be trivial, particularly if the device itself (or even just its packaging) carried a unique identifier. This could be captured and stored in the patient's notes. It was clear that this need not apply to all medical devices: it wouldn't appear to be necessary for, say, cannulas but essential for, say, stents and pacemakers.

The idea of a CONSORT type process was raised where all devices could be tracked from their manufacture and use. This would help ensure that not only devices that were used are recorded and were traceable but that devices that were not used for whatever reason were also tracked and this could provide useful data for analysis.

However, not all manufacturers necessarily have the same level of integrity.

Missing data was a concern as was data integrity, security and patient confidentiality.

It was also highlighted that for some devices it was important to record more than just the identifier for the device implanted: how the surgeon carried out the procedure could be important. For example, in inserting a stent, the balloon pressure used can affect subsequent operation so would need to be recorded contemporaneously.

There was a discussion about informed consent and how that can properly be obtained, particularly post Montgomery. This is even more important for a device with limited evidence of efficacy safety or where the patient is essentially a trial participant.

## Group C - Risk vs numbers – where do we concentrate?

### Cook A Handout

#### Risk v Numbers: Where do we concentrate?

Should attention be focussed on the most 'risky' procedures, or those where the most patients are exposed?

1. This appears to be a false dichotomy. When managing risk, a conventional (possibly overly simple) approach is to take a matrix of the probability of an event occurring, and the likely consequences if it does, then decide what and how to mitigate based on the risk of an event occurring, and the likely consequences – the consequences being a product of the effect the event produces and the number of people affected. Thus a major consequence (eg death) to few people may be more important to mitigate than a minor consequence (eg controllable bleeding associated with device insertion) to many.
2. The cells of the matrix are then coded with the risk (a product of the consequence and the likelihood) – which may be expressed numerically, or in the example below (Table 1) as green for low, yellow for medium, and red for high.

Table 1 - Example 3x3 risk matrix

		Consequences		
		Mild	Moderate	Serious
Likelihood	Low	Green	Green	Yellow
	Medium	Green	Yellow	Red
	High	Yellow	Red	Red

3. Therefore attention might be concentrated in the more leftward and lower (eg the redder) cells in the matrix.
4. However, this approach relies on being able to take a judgement on both the likelihood and consequence of an event.

How do we decide how much evidence is needed in relation to the type of device or implant being assessed?

5. Why should the approach differ from that applied to drugs? Neugebauer et al considered that there were extensive barriers to undertaking research to the same quality on devices, compared to pharmaceuticals [1]. These arguments don't seem to hold much water, as discussed in Table 2. It seems to me that first in class devices should be held to similar standards as pharmaceuticals.

Table 2- Perceived Barriers to Device Trials

Neugebauer's Argument	Problems
By the time a clinical trial is completed, a new model to	Devices change more than drugs - but can test as a class / adapt as goes along etc - there are ways to do it.

<b>Neugebauer's Argument</b>	<b>Problems</b>
replace the original device has already been developed	If change likely to hit durability / safety -> Registry. Only need a trial if doubt about efficacy.  Depending on efficacy claim, trial might not need to be very big eg hip replacement reducing pain, increasing mobility
Patients and medical professionals often have clear treatment preferences and refuse randomisation	The same happens with drugs.  There are ways to deal with this – strong unevidenced preferences can be challenged [2]  Ultimately, if a device is only available in research, patients will have to enter studies to access it [3,4].
Double blinding is difficult or impossible	Rarely impossible  The authors don't consider who needs to be masked – the patient and the person making an assessment of the outcome, but not the surgeon.  Just because something is difficult doesn't mean it shouldn't be done.
Sham surgery often poses ethical challenges, and thus restricts the choice of comparator groups	Placebo surgery is accepted by ethics committees in multiple countries, including the UK. If there is genuine equipoise there is usually not an ethical issue.  MRC/NIHR have commissioned guidance on the use of placebo surgery in research, which will be published late 2019 or early 2020.
The experience and expertise of the surgeon may have a larger impact on patient outcomes than the type of device being implanted	True, but it is understood how to deal with this, either through evaluating and mandating expertise, or through expertise based trials [5].
Defining relevant outcomes is complex	No more so than for trials of pharmaceuticals
Scientific advice and clear regulatory requirements are lacking	It's possible that views from regulators may be lacking, but the scientific issues are reasonably understood.

6. But what about 'me too' devices. In pharmaceuticals 'me too' drugs have to show efficacy and safety on their own merits, whereas generics can be marketed based on data from the original agent – provided the generic is indeed the same molecule.

7. If we treat devices like pharmaceuticals then it is likely that a new device attacking a problem through a similar mechanism with a novel (ie not tested) should be treated as a 'me too' device of the same class. Hence should demonstrate efficacy on its own behalf. This implies that version 2 of a device would need to start again in demonstration of efficacy.
8. Regulators may choose to accept efficacy evidence for a class of devices, but they don't do this for drugs.
9. How similar must a device be to one which is already marketed in order to count as a generic?

How can we balance the health and safety implications of an implant or device and the number of people who are likely to be affected?

10. For some devices, as with some drugs, there won't be enough patients to test the device for efficacy. Does the concept of an orphan device make sense here?
11. Can we pool indications for use of a device? For example, a hip prosthesis could be used for management of osteo-arthritis, and for fractured neck of femur. Is it more likely that indications can be pooled than in pharmaceuticals – the question here might be does the device adequately replace the joint rather than treat the specific indication. It may be more reasonable to pool safety data than efficacy (which is something NICE IPAC will do).
12. How often are we surprised by device failures or other safety outcomes? The recent situations with metal-on-metal hips and vaginal mesh suggest we're not good at predicting which devices will cause problems.
13. The manufacturer is potentially motivated to downplay adverse events, both to continue selling the device and to minimise pay-outs to those adversely affected.
14. Together this implies that unless there is an extensive history with very similar devices, we can't predict which devices will be safe and effective beyond data in existing research.

Can we identify when there may be a sufficiently large potential for benefit or harm so that we should be pressing for some kind of monitoring or testing?

15. Do we need to? It seems that there should be (at a minimum) registries for all implantable devices (less arguable for non-implantable) with ongoing data capture and analysis.
16. This is potentially very expensive. But it appears to be unlikely that we are able to predict low risk devices.
17. There is a potential need for health economic work to establish boundaries of which devices should be registered and followed up prospectively – the reason for not doing so would be the cost outweighing the benefit. Given the cost is likely to come from the NHS (either directly, or through the NHS paying extra to manufacturers to run such schemes), then a NICE QALY threshold may be appropriate.
18. Some manufacturers do keep registries, but while they often have fair short term safety data they can fall down when measuring efficacy, or safety beyond the short term. Many

of these registries fail to meet NICE criteria for recommendation in guidelines, see Table 3 [3].

1. They only record procedures using devices for a single manufacturer
2. They don't record counter-factuals
3. They often don't have adequate access to data for 3<sup>rd</sup> parties

What thinking and debate on this topic has been taking place inside HTA and elsewhere?

National Institute for Health and Care Excellence Interventional Procedures programme

19. NICE IPAC evaluates procedures, therefore they lump similar devices together (eg implantable defibrillators from different manufacturers).
20. They rely on regulators to spot problems with individual devices and remove them from the market.
21. Where evidence comes significantly or wholly from a single device, or a limited set of what's available on the market, guidance states this but still applies to the procedure not the exact device used.
22. Encourages the use of registries for almost every device associated with a procedure, but in many cases no appropriate registries are available (see Table 3).
23. IPAC does consider 'generations' of device. Often evidence available relates to a version of device which is no longer marketed. In this case if the older version has adequate evidence, IPAC usually issues 'special arrangements' guidance, where people doing procedures using the new version of the device are required to undertake prospective data collection and audit, and tell patients that the device used has a weak evidence base [3].

Table 3- NICE's Criteria for endorsing registries

Standard	Criteria
All known procedures (all devices), without exception, are recorded in the database	Raw anonymised data available for secondary analysis and validation. Denominator data available to assess data coverage, such as sales figures and routine health service information.
The data recorded address relevant efficacy and safety outcomes and important patient characteristics	Medicines and Healthcare products Regulatory Agency/NICE and professional representatives involved in dataset design and agree final protocol. Data include details of modifications or evolution of procedure/device and numbers done for the original indication (and respective outcomes).

Standard	Criteria
Independent oversight	Independent steering group responsible for design, data monitoring and analysis. Register recorded on national database of registers. Explicit intent to publish results whatever the outcome. Process for data collection, storage and analysis independent of any particular company or any commercial interest.

National Institute for Health Research

24. The Health Technology Assessment (HTA) programme works in a similar way to NICE. The expectation for funded trials is stable technology (ie IDEAL stage 3[6]). When there is a single manufacturer trial it's usually for organisational or financial reasons - eg donation of devices. HTA's trials are usually pragmatic [7,8], and would rarely test device A v B where purported mechanisms are the same
25. The Efficacy and Mechanism Evaluation (EME) programme is more likely to fund explanatory trials [7,8], with explicit use of single manufacturer devices. Interventions tested by EME usually have weak evidence of efficacy before the study commences. EME would consider device A v B - especially in early stage adaptive designs, or where comparing A & B might throw interesting light on mechanism.

Andrew Cook  
 June 2019

## References

- 1 Neugebauer EAM, Rath A, Antoine S-L, *et al.* Specific barriers to the conduct of randomised clinical trials on medical devices. *Trials* 2017;**18**:427. doi:10.1186/s13063-017-2168-0
- 2 Donovan JL, Rooshenas L, Jepson M, *et al.* Optimising recruitment and informed consent in randomised controlled trials: the development and implementation of the Quintet Recruitment Intervention (QRI). *Trials* 2016;**17**:283. doi:10.1186/s13063-016-1391-4
- 3 National Institute for Health and Care Excellence. *Interventional Procedures Programme Manual*. London: : National Institute for Health and Care Excellence (NICE) 2016.
- 4 Plumb J, Campbell WB, Lyratzopoulos G. How guidance on the use of interventional procedures is produced in different countries: an international survey. *International journal of technology Assessment in Health Care* 2009;**25**:124–33. doi:10.1017/S0266462309090175
- 5 Beard D, Price A, Cook JA, *et al.* Total or Partial Knee Arthroplasty Trial - TOPKAT: study protocol for a randomised controlled trial. *Trials* 2013;**14**:292. doi:10.1186/1745-6215-14-292
- 6 McCulloch P, Altman DG, Campbell WB, *et al.* No surgical innovation without evaluation: the IDEAL recommendations. 2009;**374**:1105–12. doi:10.1016/S0140-6736(09)61116-8
- 7 Ford I, Norrie J. Pragmatic Trials. *N Engl J Med* 2016;**375**:454–63. doi:10.1056/NEJMra1510059
- 8 Loudon K, Treweek S, Sullivan F, *et al.* The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;**350**:h2147–7. doi:10.1136/bmj.h2147

## Group D - How can the healthcare community better support regulation for patient benefit?

McGuire N

### Reporter's notes

Draft notes in need of considerable editing:

HW Symposium – Group D

First session: How can the healthcare community better support regulation for patient benefit?

Presenter: Dr Neil McGuire, MHRA

The MHRA understands that there are currently in excess of 600,000 medical devices on the market, and this figure does not include software, apps, accessory devices, or newly classified devices. The MHRA is the UK competent authority. It is publicly funded and arms-length from the Dept of Health.

The MHRA operates the system for reporting and recording adverse incidents. A key responsibility is to investigate device related incidents and reports, monitor investigations, and take action to prevent or reduce the likelihood of recurrence with a market surveillance role. When an issue is reported to the MHRA they will endeavour to meet with the manufacturer promptly, in recognition that a problem with any device in this context could be life-critical, for example, a syringe-driver; or may be life-critical or potentially very harmful to certain population groups, for example, a device with higher than expected residual levels of metal could be more harmful to neonates or those with renal impairment.

Notified bodies are audited every 8 months by the MHRA and the MHRA can deauthorise the NB if necessary.

When an issue with a product is raised, the MHRA can issue restrictions, product recalls, can notify other organisations including the FDA, or can publish a notice or medical device alert which should alert to an issue but can allow a pragmatic local approach to decision making to take place over the use of items. A recent example was discussed where a device issue was identified with a product and a notice was issued by the MHRA allowing local decision-making over the product's continued use.

The manufacturer has a duty to continually monitor safety and performance of the device, and submit vigilance reports to the MHRA when required to do so and there is a process of audit.

However, the system of notification from all sources is currently 'trust-based' and a clear area of weakness is that the MHRA are reliant upon being notified of problems, and cooperative, collaborative working between all concerned parties. As it is an arms length public body there is a problem that with each change in government, key information is withheld from the MHRA and this lack of access to information due to political reasons results in significant delays. While publicly funded, it is also very under-resourced for the huge role it occupies and this is going to worsen with ever-growing numbers of devices on the market, as well as in the context of the MDR/IVDR and UK implementation.

There is also delay and poor information sharing between organisations. However a key source of information is from yellow card reports, which can be completed by anyone including individual clinicians and members of the public.

Where an academic institution, group or similar is involved at the reporting stage there is frequently a considerable delay in reporting a known or potential defect or device related incidents to the MHRA.

In many cases, this delay may be attributed to culturally ingrained conduct of withholding findings, failing to share results in advance of publication or at more formative stages of investigation, due to the individual or institutional imperative to publish. This has led to cases where adverse incidents or defects are notified to the MHRA but after a significant time has elapsed in which those collecting clinical information are doing so without early sharing, to protect their work and publish their findings. This was noted to be a real problem for the MHRA and something with HCPs and organisations could work on to address and encourage earlier clinical reporting of potential defects or incidents.

The responsiveness and behaviours of academic institutions and clinical groups are important in order to receive notification of issues as promptly as possible. Greater collaboration and earlier information sharing would be welcomed. Collaboration and improved information sharing would also be welcomed from other organisations, such as other competent authorities, but these are generally much more responsive and proactive in information sharing than the scenario of those seeking to publish their findings.

MHRA oversight, in terms of product surveillance and management, 'does not begin until patients are involved' and it can be a very difficult process for any interested bodies to obtain information at an earlier stage, as this may not be available from the NB and guidance on publication or sharing will depend on the local hospital or institution involved. There was discussion around this late stage of oversight and whether there is a role for earlier intervention to streamline and oversee relevant information, monitor and share reports, and take a greater role in preventing harmful incidents. A representative of a device manufacturer discussed the processes of registering a trial before market and how this pre-market stage would be an appropriate stage for regulatory catch and oversight to begin for continuous evaluation pre-market and clinically, improving sharing of safety data from similar devices, and leading to more prompt reporting and investigation of device issues from the point of performance studies. HRA notified simultaneously to MHRA at outset. Is there a role for HRA when mandatory registration occurs?

While the Yellow Card scheme allows anyone to report an issue, Dr McGuire noted there may be an increased need to engage patient organisations or groups to assist in reporting potential issues in a responsive, balanced way, to avoid the current state of reporting where a device campaign can be led by lone voices or single issue groups that can lack balance, and which can make it more difficult to identify prescient cases. Greater balance in this would support the work of competent authorities.

Is there also a role for improving HCP education on reporting: what are the barriers to current reporting? For example, a surgeon noting adverse results or with suspicions of an issue may not feel that they are sufficiently senior to report to the MHRA. Should the institutional/professional hierarchy be addressed in order to increase reporting from all sources?

## Second Session

### Risk

Risk reporting systems 'are not joined up.' There should be improved communication and collaboration relating to regulatory issues between relevant organisations globally. It was felt that global transfer of information is appropriate and attainable, notwithstanding the huge amount of data, IT system and funding implications this would have and the relative underfunding of the UK

system in comparison to the FDA. Mandatory reporting of issues and of study results was discussed with reference to Australia and the US, but it was not felt to be a coherent response to the problem in the UK.

One reason for under-reporting safety events in the UK is that these could be linked to local sanctions, e.g. never events, and sanctions and targets may make the under-reporting problem worse.

Any reporting system must be easy to use without excessive time or system demand, and not linked to local sanctions. At present issues are raised through various pathways, however, and many are not notified to the MHRA but may be notified to the local organisation instead. Clarifying pathways and reporting requirements, removing complexity or uncertainties, without imposing sanctions, was discussed as a priority issue.

Further reference was made to various types of risk derived from products that are not immediately thought of as highest risk such as implantable devices, but may pose considerable risk if faulty or in certain patient groups, and potentially in high numbers, e.g. devices in surgical and critical care.

### Registry/follow-up

For long-term follow-up to provide relevant information, it is important that adverse events are monitored, and all relevant information captured.

The endpoints used in registries are important but can easily miss problems with a device. For example, the National Joint Registry records the endpoint of revision surgery. What endpoints should be included for devices and how can this be standardised for use beyond the UK? A useful registry would capture endpoints that are wider than this, and those recorded should include all, if achievable, potential issues with a device. For example, relevant endpoints for the long-term evaluation of a specific permanent pacemaker should be able to include falls, head injuries, and other events that could be linked to device issues; and longer term outcomes for other devices which may be more difficult to link to the initial use of a device.

Information on modifications should be captured, however small they may seem, with less or no reliance on equivalence data for devices. Data publication and sharing of all information should be strongly encouraged from all sources and from all conducting studies. This includes manufacturers, academic institutions, and, e.g. implant retrieval centres which are not currently publishing their data on explanted prostheses.

There was discussion here over a single reporting portal and work ongoing with this, and discussion over the NB system in the EU compared with the FDA approach. There was some discussion over alternative approaches for the UK moving forward with a high burden and underfunding, including MHRA/IDEAL/NICE collaborating for procedures and devices.

There is a clear need, as well, for patient engagement, education, and consent in processes to agree sharing of results and follow-up if they are not made compulsory. Single, joined-up registries with mandatory inclusion and adequate anonymisation may assist. At present there is patchy inclusion on various registries aside from the NJR in orthopaedics. Many patients are missing from the Breast and Cosmetic Implant Registry, for example. This is both a provider level issue but can also arise from lack of patient consent, with current processes.

### Evidence

Issue with equipoise and levels of evidence in many device studies: how to support study design to encourage evidence and minimise use of equivalence data which may conceal problems with a device until much later.

Other points: increased vigilance requirements from IVD/ medical device manufacturers under the MDR/IVDR, relating to post-market surveillance and the issuing of periodic safety update reports; and requirements defining specific reporting requirements including defined time frames for notifying the MHRA of incidents and preparation of reports after incidents. How is this likely to differ in practice? Will these be sufficient?

Light touch regulation v legislative gap in line with FDA on publication of results (FDA Amendments Act with fine payable by institution for delayed reporting) Will there be increased in house production? Will there be increased risk during the MDR/IVDR transition period? With the example of a product alert raised by the MHRA, will EU issues and relevant transition periods affect patient safety by narrowing available supply for alternative products, encouraging continued use of a product that comes with an increased risk? Will the UK Responsible Person, if relevant, be a point for earlier integration of some of the issues raised, or continued role of MDSO.

Overall the discussion indicated a strong need for continuous clinical evaluation and feedback with sufficiently wide endpoints recorded in any registry so that potential issues could be detected at an earlier stage, administratively robust systems and funding should be prioritised, and increased attention at all levels where earlier notification of issues could provide valuable information.