



Neutral Citation Number: [2018] EWHC 1208 (QB)

Case No: HQ 12X01255 & others

**IN THE HIGH COURT OF JUSTICE**  
**QUEEN'S BENCH DIVISION**

Royal Courts of Justice  
Strand, London, WC2A 2LL

Date: 21 May 2018

**Before:**

**THE HONOURABLE MRS JUSTICE ANDREWS DBE**

-----  
**Between:**

**COLIN GEE**  
**and others**

**Claimants**

**- and -**

**DEPUY INTERNATIONAL LIMITED**

**Defendant**

**THE DEPUY PINNACLE METAL ON METAL**  
**HIP LITIGATION**

-----  
-----

**Robin Oppenheim QC, Hugh Preston QC, Jonathan Bertram, Marcus Pilgerstorfer,  
Louise Price and Tim Cooke-Hurle (instructed by Leigh Day) for the Claimants**  
**Alexander Antelme QC, Michael Spencer QC, David Myhill, Richard Sage and Lara  
Knight (instructed by Kennedys) for the Defendant**

Hearing dates: 16 October 2017 – 26 January 2018  
-----

## **Approved Judgment**

I direct that pursuant to CPR PD 39A para 6.1 no official shorthand note shall be taken of this Judgment and that copies of this version as handed down may be treated as authentic.

.....

THE HONOURABLE MRS JUSTICE ANDREWS DBE

# INDEX

	Paragraphs
<b>CHAPTER 1: INTRODUCTION</b>	<b>1-63</b>
1.1 AN OVERVIEW OF THE CASE AND MY CONCLUSIONS	1-20
1.2 AN ADVERSE REACTION TO METAL DEBRIS	21-63
CLINICAL DEFINITION OF ARMD	21-34
THE USE OF THE TERM “ARMD” IN HISTOPATHOLOGY	35-53
ALVAL	54-59
HISTOLOGICAL FINDINGS CONSISTENT WITH A CLINICAL	
DIAGNOSIS OF ARMD	60-63
 <b>CHAPTER 2: THE RELEVANT LEGAL FRAMEWORK</b>	 <b>64-186</b>
2.1 OBJECTIVES OF THE DIRECTIVE	64-80
2.2 THE MEANING OF “DEFECT”	81-100
2.3 THE CLAIMANTS’ PRIMARY CASE ON THE IDENTITY OF THE DEFECT	101-135
2.4 THE CLAIMANTS’ ALTERNATIVE CASE ON THE IDENTITY OF THE DEFECT	136-138
2.5. LEGALLY RELEVANT CIRCUMSTANCES	139-178
AVOIDABILITY, RISK-BENEFIT AND COST	144-167
LEARNED INTERMEDIARIES	168-169
REGULATIONS AND STANDARDS	170-178
2.6 CAUSATION	179-186
 <b>CHAPTER 3: THE FACTUAL BACKGROUND</b>	 <b>187-242</b>
3.1 A SHORT HISTORY OF MODERN HIP PROSTHESES	187-204
3.2 THE DEVELOPMENT OF THE PINNACLE ULTAMET	205-222
WARNINGS AND TECHNICAL INFORMATION	213-216
REGULATORY APPROVAL OF THE PINNACLE SYSTEM	217-222

<b>3.3</b>	<b>THE IMPACT OF THE INTRODUCTION OF THE NEW ARTICULATIONS</b>	<b>223-225</b>
<b>3.4</b>	<b>EVENTS LEADING TO THE WITHDRAWAL OF THE PINNACLE ULTAMET FROM THE MARKET</b>	<b>226-242</b>
	<b>CHAPTER 4: THE CLAIMANTS' CASE</b>	<b>243-500</b>
<b>4.1</b>	<b>REVISION RATES AS AN OUTCOME MEASURE OF SAFETY</b>	<b>245-289</b>
	THE USE OF A 10 -YEAR PERIOD TO MEASURE THE RATES OF REVISION	251-258
	THE EXPERT EVIDENCE ON STATISTICS AND EPIDEMIOLOGY	259-268
	USE OF HINDSIGHT TO DENOTE THE ENTITLED EXPECTATION OF SAFETY	269-274
	MATERIALITY	275-279
	THE RELIABILITY OF CRR AS A MEASURE OF SURVIVORSHIP	280-289
<b>4.2</b>	<b>THE APPROPRIATE COMPARATOR</b>	<b>290-320</b>
	THE INTRA-PINNACLE COMPARISON	293-299
	THE APPROPRIATE EXTERNAL COMPARATOR	300-317
	THE CRR OF THE EXTERNAL COMPARATOR	318-320
<b>4.3</b>	<b>THE SWEDISH HIP ARTHROPLASTY DATA</b>	<b>321-374</b>
	THE SHAR 2000 REPORT	325-334
	THE SHAR 2002 REPORT	335-342
	THE 2002 MALCHAU PAPER	343-344
	THE SHAR 2014 REPORT	345-359
	COMPARISON OF THE SWEDISH DATA WITH THE NJR DATA ON ULTAMET	360-374
<b>4.4</b>	<b>THE NATIONAL JOINT REGISTRY DATA</b>	<b>375-455</b>
	THE CRR OF THE COMPARATOR PROSTHESES	393-402
	THE CRR OF THE PINNACLE ULTAMET	403
	EFFECT OF FINDINGS OF OTHER REGISTRIES AND STUDIES	404-413

COMPARISON OF THE CRR FOR THE ULTAMET AND COMPARATOR PROSTHESES	414-416
CONFOUNDING FACTORS	417-454
Age and sex	419
Body Mass Index	420
Activity	421-430
Asymmetric Surveillance	431-432
The increase in revision rates in 2010	433-445
The impact of “outlier” surgeons	446-454
CONCLUSIONS ON THE NJR STATISTICS	455
4.5 THE NICE GUIDANCE	456-470
4.6 THE ENGINEERING ISSUES	471-483
4.7 WAS THE PRODUCT DEFECTIVE?	484-500
CHAPTER 5: THE SIX LEAD CLAIMS	501-762
5.1 IAN HALEY	503-523
5.2 DIANE EMERY	524-545
5.3 DAWN BLAKE	546-608
5.4 SYBIL STALKER	609-682
5.5 PATRICIA GARRATT	683-724
5.6 PETER WOODS	725-762

**Mrs Justice Andrews:**

## **CHAPTER 1: INTRODUCTION**

### **1.1 AN OVERVIEW OF THE CASE AND MY CONCLUSIONS**

1. The human hip is a ball and socket joint. Over time, it may become damaged and require replacement in consequence of natural wear and tear, which may be exacerbated and escalated by certain medical conditions, the most prevalent being osteoarthritis. Total hip arthroplasty, the reconstruction of the natural hip joint with an artificial prosthesis, has become a common operation. Approximately 100,000 such operations are performed annually in the UK. A successful total hip arthroplasty can make a huge difference to the patient's quality of life by improving their mobility and relieving them from debilitating pain.
2. No surgery is without risk, but the risks associated with a total hip arthroplasty are relatively small. It carries the same general risks as any other form of major invasive surgery. Hip-specific risks include infection, dislocation, leg length inequality, nerve injury, and femoral fracture. A small proportion of patients continue to report hip symptoms even after apparently satisfactory replacement surgery. Nevertheless, it is generally considered to be a very safe, durable and reliable operation.
3. In simple terms, all prostheses used in this procedure will consist of a femoral stem, made of metal, fitted into the centre of the femur (thigh bone), to which a ball-shaped replacement femoral head is attached; and a cup which replaces the damaged acetabular socket in the pelvis. Prostheses have a smaller head and acetabulum than the natural hip. The "bearing" in a total hip arthroplasty refers to the articulation between the femoral head component and the acetabular component. The latter will either be a monobloc cup, which articulates against the femoral head, or a modular cup with a liner inserted into it, whose surface articulates against the femoral head.
4. A total hip arthroplasty is to be distinguished from a resurfacing arthroplasty, a procedure by which the native femoral head is retained, but reduced and re-shaped to receive a metal component which "resurfaces" it to produce a new integrated femoral head, which articulates against a monobloc acetabular cup. This litigation is not concerned with devices used in resurfacing arthroplasties.
5. In a total hip arthroplasty, the surgeon has a choice of bearing surface combinations. The femoral head will be made of metal or ceramic material; the liner may be made of polyethylene, ceramic or metal. Of these, polyethylene is the softest material, and therefore most prone to wear; but in recent years scientists have developed tougher and more durable forms of material by cross-linking the polyethylene with gamma radiation to produce cross-linked, and subsequently highly cross-linked polyethylene ("HXLPE"). The various combinations of articulation are:

A metal head with a liner made of polyethylene (conventional or, later, cross-linked or highly cross-linked) (MoP);

A metal head with a metal liner (MoM);

A metal head with a ceramic liner (MoC);

A ceramic head with a polyethylene liner (CoP);

A ceramic head with a ceramic liner (CoC).

6. The original types of hip replacement prostheses were monobloc systems which were fixed in position by using bone cement. Over time, the product designers and manufacturers developed modular prostheses with separate cups and liners, and modular stems, which would allow the surgeon to better reproduce each individual patient's anatomy. Modular stems and heads were well established components by the 1990s, and it was also possible to have a cup and liner combination, although the facility to use different types of liner with the same cup was not developed until around the late 1990s/early 2000s.
7. The designers and manufacturers also developed uncemented implants, which became increasingly popular. The designs vary, but the aim is for the implant to be press-fitted into the bone, allowing biological fixation over time, which, if achieved, can usually be considered lifelong. It is now possible for surgeons to use a combination of fixation techniques and to use a cemented stem in conjunction with an uncemented cup ("hybrid fixation"), or vice versa ("reverse hybrid fixation"), though the latter combination is relatively rare.
8. A surgeon will select the implant that he considers is most suited to his patient, based on the balance of how much wear it might be expected to suffer, and the demands that are likely to be placed upon it in terms of the stability of the construct and its potential risk of dislocating. Dislocation is a painful event which reduces the patient's confidence and may lead to a need for revision surgery – that is, surgery to replace all or some of the components used in the primary operation. It can cause damage to surrounding tissues with an ongoing impact on function, giving rise to a risk of further dislocation. The risk of dislocation is related to head size; the smaller the components, the greater the risk. The advantage of using a large femoral head (one with a diameter of at least 36mm) is that it brings increased stability to the joint by increasing the distance that the head must lift out in order to dislocate.
9. No artificial hip prosthesis will last forever, and if a patient lives long enough, they may need to undergo revision surgery. This is generally associated with an increased risk of the complications one might expect in primary surgery, and worse outcomes. Consequently, as the expert orthopaedic surgeon called by the claimants on generic orthopaedic issues, Mr Duncan Whitwell, put it: "*ideally, you want to make the implant outlive the patient.*" However, the actual incidence of negative outcomes will depend on many factors, including the nature and extent of the revision, and the indication for revision in the individual case. Better outcomes are generally achieved in patients who have been post-operatively monitored in a way which allows for the early identification of issues necessitating revision surgery.
10. Factors which will have a bearing on the success of revision surgery will include the surgeon and surgical factors; factors relating to the patient (including their age, weight, medical history, activity levels and types of activity); the articular bearing; the fixation, design and materials of the primary hip device; the extent of any bone damage, including fractures; the degree of any soft tissue damage and associated neurological deficit; and the degree of complexity of the revision operation. Exchanging the femoral head and/or liner is more straightforward than removing the entire prosthesis. A longer and more complex procedure is associated with a greater risk of infection and may increase the need for an early re-revision. The results of revision surgery generally are superior and complication rates are lower when the

surgery is carried out in centres and by surgeons who perform greater numbers of these procedures.

11. The Defendant (“DePuy”) is the manufacturer of a hip prosthesis system for use in total hip arthroplasty, known as the Pinnacle Acetabular Cup System (“the Pinnacle System”) which was first introduced in the UK in 2002. The Pinnacle system is an uncemented modular system, within which different articulating surface combinations were available. Surgeons could select the materials from which the femoral head and liners to the acetabular cup were made, to produce the combination of materials which they thought best suited the patient. The liner materials were conventional and, later, cross-linked and highly cross-linked polyethylene, metal, and ceramic. The metal acetabular liner in the Pinnacle system was called “Ultamet”. The surgeon also had a choice of stems produced by DePuy, of which the “Corail” stem was the most popular in the UK.
12. The claims in this Group Litigation are brought against DePuy by 312 individual claimants (in respect of 341 hips) who were implanted with one or more Pinnacle prostheses where both the acetabular liner and the femoral head were made of a metal alloy, a mixture of cobalt (Co) chromium (Cr) and molybdenum (Mo), giving a MoM articulation. In each case the femoral head was 36mm in diameter (save for four cases in which it was 40mm in diameter). All the claimants claim to have suffered an adverse reaction to metal wear debris generated by their prostheses, (“ARMD”), necessitating revision surgery. In this judgment, when I refer to the “Pinnacle Ultamet prosthesis” or “the product”, I mean a prosthesis of the type described above.
13. For most of the claimants, the Pinnacle Ultamet prosthesis was implanted during primary surgery; some of them had total hip arthroplasties on both sides. The claims are being case managed together pursuant to a Group Litigation Order (“GLO”) which was approved on 31 July 2014. There are six lead claims selected from among the claimants on the group register, three selected by the claimants’ legal representatives and three selected by DePuy’s legal representatives (though one of the latter replaced a claim that was discontinued).
14. The claims are brought under Part 1 of the Consumer Protection Act 1987 (“the Act”), which implements in England and Wales the Product Liability Directive 1985 (85/374/EEC) (“the Directive”). The claimants’ case is that the prostheses supplied to them were defective within the meaning of section 3 of the Act (interpreted in the light of the relevant provisions of the Directive) and that this caused them personal injury for which DePuy is liable to compensate them.
15. This is the trial of a common preliminary issue, namely “*whether or not the defendant is liable to the claimant, subject to any development risk defence.*” It encompasses any issues of causation.
16. At the start of this trial, the claimants relied upon expert engineering evidence that criticised various design features of components used within the Pinnacle system. Some of these criticisms turned out to have been based on a mistaken factual premise and were withdrawn. Those that were persisted in did not advance the claimants’ case. It is therefore unnecessary to burden this necessarily lengthy judgment with detailed consideration of the engineering issues, which are addressed briefly in Chapter 4 of the judgment at section 4.6.

17. Most, if not all, producers of hip prostheses manufactured MoM articulations during the 2000s. Legal proceedings have been commenced in this jurisdiction against all, or almost all, manufacturers of such prostheses. The other actions have been stayed pending the outcome of this trial. Although the findings of the court are not binding on parties to those other proceedings, it is hoped that they will provide them with guidance. For that reason, and in order to ensure, so far as possible, that the relevant legal arguments were comprehensively addressed, permission was granted for interested parties to make additional written submissions on the law.
18. Pursuant to that permission the following additional legal submissions were lodged on behalf of interested parties:
  - i) From Christopher Johnston QC and Heidi Knight on behalf of the Zimmer claimants;
  - ii) From Simeon Maskrey QC, Adam Korn and Conor Dufficy on behalf of the claimants in the Wright Medical and Biomet UK Ltd claims;
  - iii) From Oliver Campbell QC on behalf of Wright Medical Technology Inc.;
  - iv) From Jonathan Waite QC, Shaun Ferris and Jack Ferro on behalf of Zimmer GmbH and Zimmer Ltd.;
  - v) From Malcolm Sheehan QC and James Purnell on behalf of Biomet UK Ltd.;
  - vi) From Prashant Popat QC and Geraint Webb QC on behalf of Smith & Nephew Orthopaedics Ltd.

I am grateful to them, and to both teams of Counsel who appeared at the trial, for their industry and the thoroughness and clarity of their presentations.

19. The Court heard and read evidence from a wide range of experts in numerous different disciplines. Whilst most of the experts, irrespective of who called them, were mindful of their duties to the Court, I regret to say that a minority of the claimants' experts were not. Some gave the appearance of acting as advocates in the claimants' cause. Sometimes that was not entirely the expert's fault, because of the approach he had been instructed to take, but others were plainly partisan, and their reports lacked the necessary balance and impartiality. That has meant that, unfortunately, I have found their evidence unreliable, and I have placed little or no weight upon it or preferred the evidence of DePuy's experts in matters that were contentious.
20. In a case of this length and complexity it is impossible to set out all the arguments or to refer to all the evidence in the judgment, but I have tried to summarise the parties' respective positions and the main arguments and key features of the evidence that they relied on. I have taken all the evidence and all the submissions into account in reaching my conclusions. My failure to mention something does not mean that it has been overlooked. For the reasons set out in this judgment I have concluded that:
  - i) The Claimants' pleaded primary case is untenable. The inherent propensity of a MoM hip to shed metal debris through normal use, to which some patients may suffer an adverse immunological reaction, is not a "defect" in the product



within the meaning of the Act and the Directive. It did not become a “defect” by reason of the recorded incidence of such adverse reactions or the calculated risk of the probability of the revision of the prosthesis on account of them.

- ii) On their alternative case, the Claimants have failed to prove that the Pinnacle Ultamet prosthesis did not meet the level of safety that the public generally were entitled to expect at the time when it entered the market in 2002. The Court was unable to conclude on the balance of probabilities that there was a materially greater risk of a Pinnacle Ultamet prosthesis failing within the first 10 years after implant than a comparator prosthesis, and thus that the product carried with it an “abnormal risk” of damage.
- iii) Accordingly, DePuy is not liable to the claimants.

## **1.2 AN ADVERSE REACTION TO METAL DEBRIS**

### **CLINICAL DEFINITION OF “ARMD”**

- 21. In order to understand the claimants’ case, and the evidence in the lead claims, it is necessary to explain what is meant by “ARMD”. This is an expression which is used by orthopaedic surgeons making a clinical diagnosis, based on all relevant available evidence prior to, or at the time of revision surgery. The evidence will include a patient’s symptoms and presentation, radiological evidence (chiefly, diagnostic ultrasound and/or Metal Artefact Reduction (“MARS”) MRI scans), and, of course, what the surgeon witnesses at the time of the revision operation.
- 22. The generic orthopaedic experts (both of whom, happily, were fair and impartial) agreed that “ARMD” is an umbrella term that is used to describe changes that can occur in the soft tissues immediately adjacent to the implant (and less commonly, to the local bony tissues also). However, as with the term “ARPD” (adverse reaction to particulate debris), it is not used consistently among orthopaedic surgeons, and the inconsistent use of this and other terminology is a factor I have had to bear in mind when reviewing the relevant literature. DePuy’s expert, Professor Hemant Pandit, rightly stressed that the term “ARMD” includes the word “adverse”, in the sense of harmful, or damaging, and Mr Whitwell agreed in cross-examination that the inclusion of the word “adverse” was important. The claimants accept that not all periprosthetic tissue changes can properly be described as “adverse” from a clinical perspective.
- 23. All artificial hip prostheses, irrespective of the materials used, will produce particulate debris. This will not always be visible to the naked eye, or even under a microscope, as some particles will be sub-micron in size (“nano particles”). The debris is mainly the product of wear, and therefore most of it will be generated at the bearing, although some wear debris may also be produced at the junctions of modular components. Debris may also be produced in consequence of the natural degradation of the materials used over time, or their reaction to fluids or chemicals in the body; for example, metal may corrode, plastics will oxidise.

24. The nature of the debris will depend on the nature of the articulating surfaces. A MoP articulation will produce polyethylene debris, although it may also produce some metal debris. A MoM articulation will obviously only produce metal debris. The amount of wear debris produced, even by the same design of prosthesis, will vary from patient to patient, and will depend on such variable factors as how the implant has been fixed by the surgeon, the nature and amount of activity undertaken, and physical characteristics of the patient.
25. All patients will react to foreign materials in the body, but the nature and extent of the reaction to debris from a prosthesis may vary significantly from person to person. The fact that there is a biological reaction to particulate debris does not necessarily mean that the reaction is clinically adverse. The body's response to an implant may be beneficial, such as the reaction of bone to grow around a component to secure it (osseointegration), or it may be neutral in terms of benefit or disadvantage. However, some reactions will cause bony and/or soft tissue destruction and aseptic loosening of the prosthesis, necessitating revision surgery. Occasionally there may be damage to the nerves and blood cells, and in extreme cases, muscle damage.
26. The immunological reaction to particulate debris may also lead to the formation of pseudotumours. In this specific context, that expression refers to the formation of a tumour-like mass in the tissues around an implant that is non-neoplastic and non-infective. Pseudotumours communicate with the hip joint and may be associated with soft tissue and/or bony periprosthetic damage. They may be solid or filled with fluid. They can present with pain or be asymptomatic. They may, or may not, have adverse clinical consequences for the patient; this may depend on the size and precise location of the pseudotumour. They can increase or decrease in size over time.
27. A fluid collection that does not communicate with the hip joint is not a pseudotumour (though it may look similar). A radiologist should be able to see the communication on cross-sectional imaging, especially if they are looking for one, although it was common ground between the two radiology experts, Dr Simon Ostlere (for the claimants) and Dr David Wilson (for DePuy) that sometimes even the most assiduous radiologist may not be able to discern a communication that is very thin.
28. If a non-communicating collection amasses in the area of the greater trochanter, it may be due to a condition known as trochanteric bursitis. This is a degenerative thickening of the bursa, usually with a small amount of fluid, related to gluteal tendon problems. Dr Wilson, who has published papers on this condition, told the Court that in his experience, oedema (swelling) in the tendons is often misdiagnosed as trochanteric bursitis. He and his colleagues see cases of trochanteric bursitis most months in their specialist unit, but not with the frequency that it is diagnosed.
29. If there *is* bursitis, the collection that occurs may be anything from a thin sliver a millimetre in diameter over an area up to the size of the bursa, (which varies between individuals, but is typically around 6cm long) to a collection of larger dimensions. As Dr Wilson explained, if the collection has become chronic and is there for a long time, it might expand and stretch out like a balloon. If the patient has a haemorrhage or infection in that area, it can increase in size and extend and stretch out the normal cavities in that area to something much larger.

30. For the purposes of this litigation, in order for a claimant to have suffered ARMD, it must be established that sufficient damage was caused to the soft tissues by the biological reaction to metal debris from their Pinnacle Ultamet prosthesis to be regarded as clinically adverse, with or without the formation of a pseudotumour.
31. Unfortunately, there is no agreed set of diagnostic criteria for this condition, and orthopaedic surgeons have varied in their approaches to the diagnosis: this variety is illustrated by the evidence in the six lead claims. In diagnosing ARMD, some surgeons have relied on features or presentations which are neutral or non-specific, or for which there is no scientific evidence that their presence has anything to do with ARMD, such as, for example, the colour or opacity of fluid drawn from a MoM hip joint, or the presence of metal staining in the tissues (“metallosis”). As Mr Whitwell readily accepted, the presence of metallosis does not tell the clinician one way or the other whether there has been a reaction to the metal, let alone whether any such reaction is adverse. Although, in this judgment, the term “metallosis” is used in the above sense, it is yet another term which can cause confusion: some clinicians used it as a synonym for ARMD.
32. Likewise, some orthopaedic surgeons have sought to draw clinical conclusions from the level of metal ions to be found in a patient’s blood. Small amounts of Co and Cr are to be found in everyone’s blood. One would naturally expect someone with a metal implant to have a higher amount of metal ions in their blood than someone who did not have such an implant. There is much ongoing scientific debate as to what constitutes a “normal” level of ions in such a patient. However, it was common ground among the experts that there is no direct correlation between the level of metal ions in a patient’s blood, and the existence or extent of any reaction to metallic debris.
33. Although, as I will explain later in this judgment, the levels of Co and Cr ions in the blood have been used as a screening tool to help clinicians to identify patients who might have ARMD, they could not be used (and should not have been used) as a diagnostic of that condition, because a patient with higher levels than the screening threshold might have no problems at all, whereas another patient with lower levels of metal ions in the blood might suffer an adverse reaction. Mr Whitwell confirmed this from his own clinical experience.
34. The propensity of a given patient to suffer an adverse clinical reaction to particulate debris is unpredictable, and despite ongoing research, scientists have not yet worked out why one patient’s body can tolerate a particular level of debris or its products (such as a given percentage of metal ions in the blood) without suffering any adverse effects, whereas another will suffer a very serious adverse reaction with the same, or lower, levels of debris in the body.

#### **THE USE OF THE TERM “ARMD” IN HISTOPATHOLOGY**

35. “ARMD” is also an expression which is used by histopathologists as a description of tissue changes that they can see macroscopically or microscopically. Since what they are describing is a biological reaction to foreign particles, in terms of inflammation and repair, and they cannot determine whether it is clinically “adverse” or reasonably well tolerated by the patient, the use of the expression by histopathologists needs to be treated with caution. Histopathologists may refer to “*an* ARMD” rather than simply “ARMD” as a means of distinguishing between what they are describing and a

clinical diagnosis (a subtle distinction which caused considerable confusion when the claimants' expert histopathologist in the lead claims, Professor Freemont, gave his evidence).

36. Since biopsies are very rarely, if ever, taken prior to revision surgery, the histopathology evidence will not be available until afterwards, when samples of excised tissue have been examined, and therefore can only serve to confirm (or not) a clinical diagnosis that has already been made, based on other evidence. A histopathologist cannot determine whether a patient has suffered from ARMD, in the clinical sense, but he can explain whether the macroscopic and microscopic findings are consistent with that diagnosis, and identify other possible explanations for those findings. In a case where a clinician has diagnosed ARMD, one would generally expect the histopathology findings to be consistent with that diagnosis. If they are not – for example, if there is no visible soft tissue damage - then in the absence of a plausible explanation for the inconsistency (e.g. tissue sampling error, or a mistake by the histopathologist), the histopathology findings would be a strong indication that the diagnosis was incorrect.
37. The Court had the advantage of hearing evidence from two distinguished experts on the generic histopathology issues in this case, Professor Nick Athanasou from the University of Oxford, on behalf of the claimants, and Professor Scott Nelson, from the University of California, Los Angeles, on behalf of DePuy. Professor Nelson also acted as the defence histopathology expert in the six lead cases. Both were impressive, impartial and careful witnesses who were doing their best to assist the Court, and mindful of their duties as experts. This was reflected in the fact that much of their generic evidence was uncontroversial, although, as one might expect, there were differences in emphasis.
38. There were one or two passages in Professor Athanasou's expert report (such as paragraph 3.3) that were open to misunderstanding because they did not contain all the detail necessary to paint a full and balanced picture, but I am sure, having seen and heard him, that this was due to his trying to simplify a complex subject in a way that a lay person could easily follow. Although Professor Athanasou included in his report some observations about the prevalence and progression of pseudotumours in different bearing surfaces, I place no weight on those observations in the light of his fair acceptance in cross-examination that he was not really qualified to comment on the papers on that subject to which he referred, and that he had not even read through them all. Subject to those caveats, I found Professor Athanasou's evidence helpful. The following simplified explanation of the biological reaction to foreign debris is based upon his and Dr Nelson's evidence and on that of the two (similarly impressive) expert immunologists, Professor John Kirby and Professor Ian Kimber.
39. The human body will produce immune responses to wear debris from any form of prosthetic implant. The immunological reactions that debris can produce in a patient's body are fundamentally similar, although the specific reactions, outcomes and clinical consequences may vary in proportions between the different types of prostheses. The actual types of reaction and the actual consequences, such as soft tissue damage or bony damage, or the formation of pseudotumours, can occur across the entire range of prostheses.

40. There are two categories of immune responses, innate and adaptive. “Innate immunity” is a more primitive, non-specific, immune system, that by a variety of means provides a first line of host defence before the (secondary) adaptive immune response may become engaged. The innate immune system plays an important role in supporting adaptive immune responses. It is a complex response in which many factors come into play. “Adaptive immunity” is a sophisticated and dedicated immune response that has the cardinal features of memory, specificity and the ability to distinguish between self and non-self, and to mount specific immune responses. There cannot be an adaptive immune response unless there has first been an innate immune response.
41. All forms of foreign debris can trigger an innate immune response (although potentially involving different mechanisms). Macrophages play an important part in that response. They are cells which may be activated by a wide variety of stimuli, and they serve several important functions. Their presence is a signal of chronic inflammation, and therefore macrophages will be present to some extent in a patient who is already suffering from a chronic inflammatory disease such as osteoarthritis, irrespective of whether the patient has had a joint replacement. They are responsive to, and also produce, chemical mediators such as cytokines (proteins that, by their effect on lymphocytes and/or other cells regulate the nature, intensity and duration of the immune response).
42. The products of activated macrophages attempt to eliminate perceived injurious agents or materials from the body, be they indigenous or foreign in nature. They do this by bringing the material into the macrophage cell and breaking it down (a process known as phagocytosis), so that in due course the harmful elements can be secreted from the body. If one cell cannot achieve this, for example because the particle is too large, it will recruit others. If need be, macrophages can fuse with each other to form multinucleated foreign body giant cells.
43. Macrophages also play an important role in the adaptive immune response. In some individuals, macrophages will ingest and present antigens to T-cells (T lymphocytes) which will activate those cells and produce what scientists call type IV hypersensitivity, or what a lay person would describe as an allergic reaction. It is potentially a much more complex reaction than with innate immunity, given the number of processes involved, and again will vary from person to person.
44. An adaptive response is not necessarily more damaging or longer lasting than the innate response – it may be, but it depends on the individual. Some people are predisposed to suffer such a reaction, others may become sensitised over time. The experts agreed that there is no reliable method for identifying those who are predisposed to mounting a cell-mediated hypersensitivity reaction to a MoM implant. Professor Kirby said there had been less research into the potential in humans for an adaptive response to cobalt than for other metals, such as nickel, and virtually none into chromium.
45. It is clear from the evidence of Professor Athanasou and Professor Nelson that it would be a mistake to equate a cellular reaction with a clinical consequence. There is a spectrum of innate foreign body macrophage response to all particles generated from a hip prosthesis, irrespective of the material used. There will always be some non-specific inflammation present in the tissues of any patient with an implanted

device of any type, but it may not be significant in type or amount, and it may have no clinical consequences. Mr Whitwell accepted that the presence of macrophages, even in large numbers, may have no adverse clinical consequences for a patient.

46. If wear particles cannot be completely broken down, they may remain in the macrophage (or giant cell) for the natural life of the cell. They may do so without causing any harm. However, in some cases the phagocytosed macrophages will release chemicals that can promote inflammation and osteolysis (the destruction of bone) by stimulating the formation of osteoclasts (cells whose sole function is to cause bone resorption) and consequential resorption of periprosthetic bone, eventually resulting in aseptic loosening of the prosthesis. This is agreed by the experts to be a more common reaction to polyethylene debris than to metal debris. In other cases, the particle size, shape and/or chemical composition of the wear particles, or their products when broken down, may prove cytotoxic, that is, cause or accelerate the death of the phagocytosed macrophages and other cells such as lymphocytes, which may in turn lead to soft tissue necrosis.
47. It is no part of the claimants' case that metal debris gives rise to systemic effects. Activation of the innate immune response, by whatever trigger, may lead to inflammation, but not necessarily to lasting soft tissue damage. Even tissue that is damaged in consequence of an immune reaction may be naturally repaired as part of the biological process of healing; dead cells may be cleaned up and replaced by others that function normally. Thus, even if there is a cytotoxic reaction to debris (of whatever sort), the death of cells may not lead to an adverse clinical consequence in the form of the death of tissue. Tissue is not just a collection of cells; it contains blood vessels, collagen, and other components. When there is tissue necrosis, all of these will be dead. Professor Nelson described it thus: *"if you have a piece of tissue and you see a lot of dead cells and really nothing in between, that is tissue necrosis"*.
48. All cells, including macrophages, have a limited lifespan and will die naturally. In theory there would be nothing alarming about seeing a few dead or dying cells appearing in tissue examined under the microscope, but as Professor Athanasou observed, in practice it is well-nigh impossible to pick out a single dead cell on a slide. The presence of significant quantities of dead cells in tissue is a sign of pathology and a cause for concern, because it suggests that something is causing those cells to die. Sometimes the histopathologist can discern a layer of dead cells, or dying cells, next to a layer of normal cells, which may suggest a progressive necrotic process within that tissue. However, a histopathologist can only speak to the state of the tissue he sees on the slide at the time of examination. He may express a view as to what may have caused it and over what period, but how that situation may evolve in the future is a matter for a clinician.
49. Professor Nelson and Professor Athanasou agreed that the death of macrophages induced by metal or polymer particles is size and concentration dependent. Indeed, there was general consensus across the various relevant scientific disciplines, including mechanical engineering, that the volume dose and rate of generation of wear and wear debris produced by a prosthesis has an impact on its potential for inducing adverse biological reactions. In very broad terms, the greater the dose, and the greater the concentration in the tissue, the more likely there is to be a cytotoxic reaction (and vice versa), though tolerance levels will vary from patient to patient.

50. The required dose to trigger a cytotoxic reaction to metal debris has not yet been estimated with any degree of accuracy. Experiments have been carried out and scientific research is ongoing, but it is extremely difficult, if not impossible, to replicate in a laboratory what goes on in the human body. I heard evidence on that topic from Professor John Fisher, a professor of Mechanical Engineering at the University of Leeds, who is a leader in the field of mechanical engineering in the context of hip joint replacement prostheses, and from Professor Kimber, who is a toxicologist as well as an immunologist.
51. Based on scientific research that has been carried out to date, both these experts took the view that the amount of metal debris generated by a normally-functioning MoM hip prosthesis was unlikely to cause ARMD unless the patient suffered a hypersensitive reaction. However, they accepted that more research needs to be carried out before it can be reliably established whether their view is correct. Professor Fisher deferred to Professor Kimber as having more relevant expertise. Professor Kimber's evidence, which I accept, is that there is a minimum safe threshold dose beyond which there is no risk of ARMD in humans (absent exceptional hypersensitivity) but that scientific experiments have not yet discerned exactly where that line is to be drawn in mathematical terms. Professor Kimber also agreed with Professor Kirby that the biological response is dependent on the sensitivity of the individual patient, irrespective of the amount of wear in the implant.
52. So far as particle size is concerned, it was agreed by all the relevant experts that micron sized particles may drive an *innate* inflammatory response, but nano sized particles are more likely to drive an *adaptive* immune response. The likelihood of an adaptive immune response is higher with metal ions than with polyethylene or ceramic debris. With ceramic debris, the immunological response would be less extensive, and predominantly macrophage in nature.
53. One feature of the adaptive response may be an aggregation of lymphocytes around small vessels in periprosthetic tissues ("perivascular cuffing"). Cuffing is not specific to the type IV hypersensitivity reaction or to MoM bearings, but where there is a type IV hypersensitivity reaction to metal debris, the lymphoid response may be far more pronounced. In his supplemental expert report, Professor Athanasou agreed with Dr Nelson that perivascular lymphocytes and lymphoid aggregates, as well as a macrophage infiltrate, can be seen in a patient with osteoarthritis, although he said that in such a patient the macrophage infiltrate is not as marked as in ARMD. He added that the features one would expect to see in a case of ARMD were "*extensive necrosis associated with deposition of Co-Cr particles, a heavy foreign body macrophage and an ALVAL response.*"

### ALVAL

54. The term "ALVAL" was first coined by Dr Willert and others in a seminal clinical study published in the Journal of Bone and Joint Surgery in 2005, to describe a condition seen under the microscope, a lesion in which a distinct lymphocytic infiltration is present, with perivascular cuffing, accompanied by plasma cells, and visible metal debris, often inside macrophages but sometimes outside them. Other features they noted in some patients were drop-like inclusions in the cytoplasm of the macrophages (though Professor Athanasou said he has never seen these), areas of cell necrosis, often large, and extensive fibrin exudation. Fibrin is a generic term for

plasma protein and is a pink material which histopathologists see in the context of inflammation. It is commonly seen on the surface of the joint capsule in osteoarthritis and after joint replacement.

55. The orthopaedic and histopathology experts agreed that ALVAL is a purely histological finding which takes account of the various features of the tissues observed under the microscope. The one essential feature of it is the distinct diffuse perivascular lymphocytic infiltration, and the extent to which some of the other features mentioned in the Willert paper are present may contribute to a conclusion that ALVAL is present.
56. Again, there is a danger of terminological confusion because some clinicians have used the expression “ALVAL” as a synonym for ARMD (in the clinical sense). However, the fact that an ALVAL response may be seen by a histopathologist when examining tissue taken from a patient does not necessarily mean the patient has ARMD (in either sense). A significant lymphocytic infiltration *may* be consistent with the histological finding or the clinical diagnosis of ARMD, particularly if it is accompanied by significant tissue necrosis and the presence of metal debris in and around the dead or dying cells. However, Mr Whitwell accepted that the mere fact of an ALVAL reaction does not mean that there has been an *adverse* reaction in clinical terms. It depends on the nature and extent of the ALVAL response and any other relevant features.
57. It was common ground that there can be ARMD (in both the clinical and histological sense) without ALVAL, but in Professor Athanasou’s extensive experience, it was rare to find ARMD without a lymphocytic response. He said that when one did, it was associated with heavy necrosis, macrophage infiltrate, and a large pseudotumour. This evidence is of some importance when considering the histopathology in the six lead claims.
58. Professor Nelson said that tissue necrosis may be more extensive in “high ALVAL” cases. This was consistent with Professor Athanasou’s evidence and is also borne out by the histological findings in two of the lead cases, those of Mr Haley and Mrs Emery. Professor Nelson was referring to a scoring system for ALVAL that was developed by a team headed by Dr Patricia Campbell, on which he was the pathologist, by ranking the multiple features described in the Willert study. The Campbell system was originally used as part of a study of the correlation between the incidence of ALVAL and the amount of prosthetic wear measured in explants from patients who had had revisions (not just of total hip arthroplasties but of resurfacing arthroplasties). A different scoring system for ALVAL, used for a similar purpose, but focusing on the lymphocytic infiltrate, was developed in Oxford by Dr Grammatopoulos and a team including Professor Athanasou and Professor Pandit.
59. Professor Athanasou was emphatic in his evidence that neither scoring system was intended to be, or should be used as, a diagnostic tool (and was critical of those who did use it as such). He accepted that whichever system was used, there was likely to be broad agreement between histopathologists as to where in the spectrum the ALVAL phenomenon observed in the tissue lay. That has proved to be true in this trial.



## **HISTOLOGICAL FINDINGS CONSISTENT WITH A CLINICAL DIAGNOSIS OF ARMD**

60. The histopathology experts agreed that a macrophage response to metal wear particles is a relevant feature of ARMD (in the histological sense). Fibrin or a lymphocytic response (including a low-grade lymphocytic response) are non-specific to ARMD. Bone and/or soft tissue necrosis, macrophage response to wear particles, the formulation of granuloma (a distinctive pattern of chronic inflammation characterised by the presence of numerous macrophages), and lymphoid infiltrate are all histological features that can support the clinical diagnosis.
61. The experts agreed that the macrophage response probably reflects the innate response and the lymphocytic response the adaptive response. Therefore, if, in a case of ARMD, in the clinical sense, the histopathology reveals significant ALVAL, it can be deduced that the adaptive immune response and cell-mediated hypersensitivity caused or contributed to that condition. If there is ARMD without ALVAL it is most likely to be the patient's innate immune system that is driving the reaction, and the condition has probably been brought about by a quantity and concentration of metal or metal products sufficient to bring about a cytotoxic effect.
62. As to the correlation between pseudotumours and tissue necrosis, Professor Athanasou's evidence was that almost all cases in which a very large amount of necrosis is found are associated radiologically and clinically with very large pseudotumours. When asked if he would expect to see extensive tissue necrosis in a patient with ARMD, Professor Athanasou said he would expect to see that commonly in cases where there is a pseudotumour, but that it had been found in a significant number of patients without a pseudotumour. Most patients with failed implants will generally have extensive necrosis, but they may have less. He said that whilst minimal necrosis can be seen in ARMD (in the histological sense), there is a spectrum of histological findings and it is more common to see more extensive necrosis. If there were minimal necrosis, he would want to see other strong features to point towards a finding of ARMD, because one of the critical features would be missing. He said that generally, in a case of ARMD when you find a significant macrophage response, you do see evidence of necrosis of those macrophages.
63. It was agreed by the experts that the state of the synovial lining is not a histological feature that is seen in ARMD. The synovium is the tissue covering the inner part of the joint capsule. The synovial membrane lines the joint cavity except over the articular cartilage, and the membrane is covered by a synovial lining which is normally one or two cells in thickness. Beneath this there is connective tissue. Dr Nelson's evidence was that if areas of the synovial lining are missing, that may be consistent with ALVAL, but it is also seen in many other settings. Therefore, the state of the synovial lining is of little or no assistance in determining whether there is ARMD.

## **CHAPTER 2: THE RELEVANT LEGAL FRAMEWORK**

### **2.1 OBJECTIVES OF THE DIRECTIVE**

64. The legislative process leading from the initial proposal by the European Commission in 1976 to the adoption of the Directive in 1985 was a long one, and the Directive in its final form represents a compromise reconciling the different interests at stake. It

was necessary to balance the interests of consumers and producers in a manner acceptable to all Member States, and since there were initially wide differences of approach to the question of where precisely that balance should be struck, negotiations on that question took a long time.

65. A helpful summary of the evolution of the Directive was provided by Advocate General Tesouro in his Opinion of 23 January 1997 in *Commission v UK* (Case C-300/95) at [15] to [18]. He described the Commission’s original proposal as “one of absolute liability in that the producer could put forward no evidence in rebuttal” and then explained, at [19]:

*“In contrast [to the original Commission proposal] the Directive as it was adopted by the Council opted for a system of strict liability which was no longer absolute, but limited, in deference to a principle of the fair apportionment of risk between the injured person and the producer, the latter having to bear only quantifiable risks, but not development risks which are, by their nature, unquantifiable. Under the Directive, therefore, in order for the producer to be held liable for defects in the product, the injured party is required to prove the damage, the defect in the product and the causal relationship between defect and damage, but not negligence on the part of the producer.*

*The producer, however, may exonerate himself from liability by proving that the “state of the art” at the time when he put the product into circulation was not such as to cause the product to be regarded as defective. That is what Article 7(e) of the Directive provides.”*

66. Thus the final version aims to strike a fair balance between the various competing interests by introducing a system of no-fault liability for products that fail to meet the standard of safety that, in all the circumstances, the public are entitled to expect, subject only to specific defences which the producer bears the burden of establishing, including the so-called “development risk defence” under Article 7(e) that was the subject of consideration in that case. Member States are entitled to derogate from Article 7(e), but they must follow a specific process if they wish to do so.
67. The Directive is not directly applicable, and Member States were therefore left to implement it by means of their own domestic legislation. The claimants’ cause of action is therefore derived from the Act. It is well established that domestic legislation which brings into effect an EU directive must be interpreted, so far as is possible, in the light of the wording and the purpose of the directive, so as to achieve the result intended by the latter: *Marleasing SA v La Comercial Internacional de Alimentacion SA* (Case C-106/89) [1992] 1 CMLR 305. Here, that interpretive obligation has been reinforced by the express language of section 1(1) of the Act, which provides that Part I “shall have effect for the purpose of making such provision as is necessary in order to comply with the Product Liability Directive and shall be construed accordingly.” Therefore, when construing the provisions of the Act, the Court must bear closely in mind the underlying purposes of the Directive.
68. In this litigation, the claimants have not suggested that the Act fails to reflect the Directive in any material respect, but they have sought to place reliance on the Directive (especially its recitals) and on the travaux préparatoires in support of their primary case on defect. They also contend that these materials provide support for a

restrictive interpretation of the circumstances that the Court is entitled to take into consideration when evaluating whether a product is defective.

69. The purposes of EU legislation may be ascertained from its recitals, especially any objectives that are specifically referred to in those recitals. A core objective of the Directive, reflected in its first recital, was to achieve total harmonisation of strict product liability throughout the EU, irrespective of the identity of the product. Variations in the burdens imposed on producers by different national conditions for product liability were thought to create unacceptable distortions in competition and impediments to the free movement of goods within the common market. Therefore, the first justification given for the Directive is an economic one.
70. Against the background of that perceived need for harmonisation, as the second recital makes clear, the conclusion reached by the relevant legislative body – in this case, the Council of Europe - on the issue of how to strike a fair balance between the interests of consumers and producers was that:
- “... liability without fault on the part of the producer is the sole means of adequately solving the problem, peculiar to our age of increasing technicality, of a fair apportionment of the risks inherent in modern technological production.”*
71. Many of the recitals refer specifically to the protection of the consumer or the effective protection of the consumer. Thus, for example:
- i) Recital 4 states that protection of the consumer requires that all producers involved in the production process should be made liable, in so far as their finished product, component part or any raw material supplied by them was defective.
  - ii) Recital 5 states that in situations where several persons are liable for the same damage, the protection of the consumer requires that the injured person should be able to claim full compensation for the damage from any one of them.
  - iii) Recital 6 states that to protect the physical well-being and property of the consumer, the defectiveness of the product should be determined by reference not to its fitness for use but to the lack of safety which the public at large is entitled to expect, whereas the safety is assessed by excluding any misuse of the product not reasonable under the circumstances.
  - iv) Recital 12 states that to achieve effective protection of consumers, no contractual derogation should be permitted as regards the liability of the producer in relation to the injured person.
  - v) Recital 13 provides that claims for damages under national legal systems based on contractual or tortious causes of action should remain unaffected by the Directive in so far as these provisions also serve to attain the objective of effective protection of consumers.
72. However, the recitals also contain examples of the complex balancing between numerous competing policy considerations which is reflected in the substantive provisions of the Directive, and in turn reflected in the Act. Thus, for example:

- i) Recital 7 states that “a fair apportionment of risk between the injured person and the producer implies that the producer should be able to free himself from liability if he furnishes proof as to the existence of certain exonerating circumstances”. This is a reference to the defences set out in Article 7, including the development risk defence under Article 7(e).
  - ii) Recital 8 states that the protection of the consumer requires that the liability of the producer remains unaffected by acts or omissions of other persons having contributed to cause the damage, but that the contributory negligence of the injured person may be taken into account to reduce or disallow such liability.
  - iii) Recital 10 states that a uniform period of limitation for the bringing of action for compensation is in the interests both of the injured person and of the producer.
  - iv) Recital 11 provides that whereas products age in the course of time, higher safety standards are developed and the state of science and technology progresses, therefore it would not be reasonable to make the producer liable for an unlimited period for the defectiveness of his product.
73. This indicates that, whilst the effective protection of consumers is a key objective of the Directive, it is not the main or overriding objective. It has equal status with the other objectives. It is important to bear this in mind.
74. It is legitimate for the Court to have regard to the travaux préparatoires as an aid to interpretation of the Directive (and thus the Act), provided that appropriate caution is exercised to ensure, as far as possible, that any views or objectives expressed at an earlier stage of the long drawn-out legislative process were in fact adopted or reflected in the final decisions taken by the relevant legislative body, in this case, the Council of Europe. Care is called for when looking at the travaux in the present case, because the original regime proposed by the European Commission was ultimately rejected. Justifications put forward for that original proposal are not reliable aids to interpretation, especially if they were justifications for imposing a system of absolute liability on producers. The same is true of recitals to the draft Directive which did not find their way into the final version. Indeed, it can be inferred from their omission that they should not be relied upon.
75. On behalf of the claimants, Mr Pilgerstorfer took the Court through a selection of the travaux, but they added nothing of any relevance to what can be gleaned from looking at the Directive itself. At most, they reconfirmed the objectives stated in the recitals.
76. The claimants acknowledged that the Directive seeks to set a balance between the interests of producers and consumers, but they submitted that this does not tell the Court precisely where the balance has been set. I disagree. The Directive makes it clear that the balance is struck by providing for no-fault liability if a defective product causes damage, subject only to a limited number of clearly defined defences. That is why the CJEU, in three decisions in 2002, *Commission v France* [2002] ECR I-3827, *Commission v Greece* [2002] ECR I-3879 and *Gonzalez Sanchez v Medicina Asturiana SA* [2002] ECR I-3901, decided that the Directive provided for complete harmonisation, and that national legislature did not retain the power to provide

consumers with a higher level of protection in relation to liability for damage caused by defective products.

77. In each of these cases the CJEU pointed out that it was important to note that the Directive contains no provision expressly authorising the Member States to adopt or to maintain more stringent provisions in matters in respect of which it makes provision, in order to secure a *higher* level of consumer protection. The reference in Article 13 of the Directive to the rights which an injured person may rely on under the rules of the law of contractual or non-contractual liability must be interpreted as meaning that the system of rules put in place by the Directive (in particular under Article 4) does not preclude the application of other systems of contractual or non-contractual liability *based on other grounds* (than defect), such as fault, or a warranty in respect of latent defects. Thus, for example, in the *Gonzalez Sanchez* case, the CJEU decided that Article 13 must be interpreted as meaning that the rights conferred under the legislation of a Member State on the victims of damage caused by a defective product under a general system of liability having the same basis as that put in place by the Directive, may be limited or restricted as a result of the transposition of the Directive into the domestic law of that State.
78. Neither the travaux nor decisions of the CJEU relating to the Directive establish that the aim of the Directive was to provide a simple or straightforward route to redress for consumers *beyond* the introduction of no-fault liability. Statements about a “higher form of protection” for consumers must be understood in the context that a fault-based system contained more obstacles for the consumer to overcome than the system of no-fault liability introduced by the Directive, and even a reversed burden of proof was considered to afford insufficient protection. Only no-fault liability, subject to carefully circumscribed defences, would suffice. National courts were left to decide the procedure and the rules of evidence by which such liability would be established.
79. The CJEU has been vigilant to ensure that the rights conferred by the Directive are effective, irrespective of whether those rights are conferred on consumers or producers. Thus in *NW and others v Sanofi Pasteur* (Case 621/15) (2017) ECLI:EU:C:2017:484 (“*Sanofi Pasteur*”) the CJEU held that although the Directive left it to each Member State to establish detailed rules of proof and evidence for practical implementation of the Directive, which might vary depending on the type of product involved, national rules governing how evidence is to be adduced and appraised must not be such as to undermine either the apportionment of the burden of proof provided for under Article 4 or, more generally, the effectiveness of the system of liability provided for under the Directive, or the objectives pursued by the EU legislature by means of that system.
80. At [36] – [37] of its judgment the CJEU warned against national courts applying evidentiary rules in such a way that, where one or more types of factual evidence were presented together, an immediate and automatic presumption would operate of there being a defect in the product and/or a causal link between that defect and the occurrence of the damage:

*“Therefore, national courts must first ensure that the evidence adduced is sufficiently serious, specific and consistent to warrant the conclusion that, notwithstanding the evidence produced and the arguments put forward by the producer, a defect in the product appears to be the most plausible explanation for the occurrence of the*

*damage, with the result that the defect and the causal link may reasonably be considered to be established.”*

It is clear from this decision that any interpretation of a domestic statute which would operate in a way that obviated the necessity for a claimant to prove the defect, or the causal link between defect and damage, would be as much contrary to the objectives of the Directive as a provision that had the practical effect of widening the limited defences available to a producer under Article 7.

## **2.2 THE MEANING OF “DEFECT”**

81. Section 2(1) of the Act provides that:

*“where any damage is caused wholly or partly by a defect in a product, every person to whom subsection (2) below applies shall be liable for the damage.”*

Article 1 of the Directive provides that *“the producer shall be liable for damage caused by a defect in his product”* and Article 4 states that *“the injured person shall be required to prove the damage, the defect and the causal relationship between defect and damage”*. It is therefore clear that, consistently with the Directive, the Act creates a liability without fault, and that all that the claimant needs to prove is (1) that there was a defect in the product in question and (2) that the defect caused him to suffer damage. The nature of the liability imposed is unique to the Act, based on the definition of defect.

82. Section 3 of the Act defines “defect” as follows:

*(1) Subject to the following provisions of this section, there is a defect in a product for the purposes of this Part if the safety of the product is not such as persons generally are entitled to expect; and for those purposes “safety”, in relation to a product, shall include safety with respect to products comprised in that product and safety in the context of risk of damage to property, as well as in the context of risks of death or personal injury.*

*(2) In determining for the purposes of subsection (1) above what persons generally are entitled to expect in relation to a product all circumstances shall be taken into account, including –*

*(a) the manner in which, and purposes for which, the product has been marketed, its get up, the use of any mark in relation to the product and any instructions for, or warnings with respect to, doing or refraining from doing anything with or in relation to the product;*

*(b) what might reasonably be expected to be done with or in relation to the product; and*

*(c) the time when the product was supplied by its producer to another;*

*and nothing in this section shall require a defect to be inferred from the fact alone that the safety of a product which is supplied after that time is greater than the safety of the product in question.”*

83. Article 6 of the Directive provides as follows:

*"1. A product is defective when it does not provide the safety which a person is entitled to expect, taking all circumstances into account, including:*

- (a) the presentation of the product;*
- (b) the use to which it could reasonably be expected that the product would be put;*
- (c) the time when the product was put into circulation.*

*A product shall not be considered defective for the sole reason that a better product is subsequently put into circulation.”*

84. It was common ground that the level of safety that the public is entitled to expect must be evaluated at the time when the product is first put on the market by the producer, though strictly speaking, that time is one of the circumstances which the Court must take into account. However, in determining whether the product met that level of safety, the Court is entitled to have regard to everything now known about it that is relevant to that enquiry, irrespective of whether that information was available at the time it was put on the market or has come to light subsequently. That is obviously the correct approach, otherwise a claimant would never be able to establish that a product, whose lack of safety only comes to light one or two years after it was first marketed, was defective at the time of its initial circulation.

85. Section 3 of the Act and Article 6 of the Directive have subtle linguistic differences, but these are easy to reconcile in such a way as to produce consistency of interpretation in line with the objectives of the Directive. The Directive refers to “the safety which *a person* is entitled to expect” whereas the Act, reflecting the sixth recital to the Directive, refers to “the safety [that] *persons generally* are entitled to expect.” The test is an objective one. By using the words “*persons generally*,” Parliament has clearly removed any room for misunderstanding that the expression “*a person*” might have created, since in this jurisdiction, negligence liability is determined by the standards of the hypothetical reasonable person. Under the Act, the standard of safety is not measured by what such a person might reasonably expect, but by what anyone (and thus, the public at large) is *entitled* to expect in all the circumstances.

86. Whilst s.3 of the Act uses the phrase “there is a defect in a product” and Article 6 uses the phrase “a product is defective” they are describing the same thing. The CJEU put this beyond doubt in *Sanofi Pasteur* at [22], stating that:

*“the concept of “defect” within the meaning of [Articles 1 and 4 of the Directive] is indeed defined in Article 6 thereof.”*

The concept of “defect” introduced by the Directive is an autonomous one, defined in terms of failure to meet an objective standard of safety that the Court must evaluate. If it is unsafe by that standard, it is defective. The “defect” is therefore defined by reference to the condition of the product itself, i.e. the product’s failure to meet that level of safety, rather than by reference to some fault or deficiency in it, or any precise mechanism that caused the damage. Indeed, it may be impossible to determine what the mechanism was.

87. Thus, the dictionary definition or the normal understanding of the word “defect” plays no part in determining whether a product is defective (or whether there is a defect in it) within the meaning of the Directive or the Act. The Court’s focus should be on whether the product is safe, as measured by the test in s.3, rather than on whether there is a specific fault in it. Of course, there may be more than one reason why the product is unsafe: if the lack of safety is due to a combination of circumstances or features, then it is that combination which makes the product defective; but the defect will still consist of whatever it is about the character, state or condition of the product that makes it unsafe.
88. Article 7 (e) of the Directive provides that it is a defence for the producer to prove *“that the state of scientific and technical knowledge at the time when he put the product into circulation was not such as to enable the existence of the defect to be discovered.”* This is the so-called “development risk defence” that was enacted in domestic legislation in s.4 (e) of the Act, which provides that in respect of a defect in a product it shall be a defence for the producer to show:
- “that the state of scientific and technical knowledge at the relevant time was not such that a producer of products of the same description as the product in question might be expected to have discovered the defect if it had existed in his products while they were under his control.”*
89. The European Commission took issue with that formulation. They brought proceedings before the CJEU seeking a declaration that the UK had failed properly to implement the Directive: (Case C-300/95) *Commission v United Kingdom* [1997] 3 CMLR 923. The Commission argued that the defence under Article 7(e) was narrower than the defence under s.4(1)(e) and that the latter test was easier to satisfy. They also claimed that the national provision had the effect of transforming strict liability into a liability founded on negligence.
90. The CJEU disagreed. They accepted the argument of the UK that the test under Article 7(e) was objective, that the *“state of scientific and technical knowledge”* is a reference to the state of knowledge which producers of the class of the producer in question, understood in the generic sense, may objectively be expected to have, and that that was precisely the meaning of s.4(1)(e) of the Act. The court held that the wording of s.4(1)(e) of the Act did not suggest, as the Commission alleged, that the availability of the defence depends on the subjective knowledge of a producer taking reasonable care in the light of the standard precautions taken in the industrial sector in question. Moreover, there was nothing in the material before the court to suggest that the domestic courts would interpret that section in a manner inconsistent with the wording and purpose of the Directive.



91. The CJEU adopted the analysis of the Advocate General. It held, at [29], that in order to have a defence under article 7 (e) of the Directive, the producer of a defective product must prove that the objective state of scientific and technical knowledge, including the most advanced level of such knowledge, at the time the product in question was put into circulation was not such as to enable the existence of the defect to be discovered. Further, in order for the relevant scientific and technical knowledge to be successfully pleaded *against* the producer, that knowledge must have been accessible at the time when the product in question was put into circulation. On this last point, the court observed that article 7(e) of the Directive raises difficulties of interpretation which in the event of litigation the national courts will have to resolve, if necessary having recourse to Article 177 EC.
92. Although at this stage of the proceedings I am not concerned with the merits of any development risk defence raised by DePuy, s.4(1)(e) is relevant, to the extent that the word “defect” must be interpreted consistently in that section and the earlier sections of the Act in which it appears. Whilst that defence should not be interpreted in a manner that would re-introduce the need for proof of fault by the back door, it is equally important that the Act should not be interpreted in a manner which unjustifiably circumscribes the defence, to the detriment of the producer.
93. The standard of safety which the public is entitled to expect from each product at the time it enters the market is a matter of law, and it may differ from product to product. The Act and the Directive apply to virtually all products supplied to consumers (every “moveable” which is commercially supplied) – from household appliances to cars, from a bottle top to a medicine. It would be impossible to set a single yardstick for safety across such a wide range. The Council plainly wished to give the national courts sufficient flexibility to be able to address the issue of defect on a case by case basis, by reference to the circumstances pertaining to any one of a diverse range of products with diverse uses and characteristics, giving rise to diverse risks. There is no distinction drawn in the Directive, or in the Act, between different types of product – “standard” or “non-standard”, for example, or between different types of defect – such as a manufacturing defect, a design defect or a warning defect, as in the US system. The definition of “defect” applies across the board, and the sole issue for the Court in determining whether a product is defective or not is whether it meets the standard of safety set out in s.3 of the Act.
94. What the public is *entitled* to expect may not match a person’s actual expectation. As Burton J observed in *A v National Blood Authority* [2001] 3 All ER 289, (“*A v NBA*”) at [31]:

*“the court decides what the public is entitled to expect... such objectively assessed... expectation may accord with actual expectation; but it may be more than the public actually expects, thus imposing a higher standard of safety, or it may be less than the public actually expects. Alternatively, the public may have no actual expectation – e.g. in relation to a new product.”* (emphasis in the original).

The Court is given the task of determining what that standard is, by reference to “all the circumstances.” The parties accepted that “all the circumstances” means “all the relevant circumstances” but there was considerable disagreement as to what type of circumstances will be legally relevant. I consider the rival arguments in Chapter 2 of this judgment, section 2.5, under the heading “legally relevant circumstances”.

95. In *A v NBA* it was common ground that the phrase “*entitled to expect*” should mean what “*the legitimate expectation is of persons generally, i.e. what is legitimately to be expected arrived at objectively.*” Burton J adopted the formulation “legitimate expectation” in his judgment: a similar formulation can be found occasionally in the European cases, and Burton J said it was analogous to the formulation in other languages in which the Directive is published. Whilst it is understandable why, in those circumstances, Burton J adopted it as a convenient form of shorthand, I agree with the observations of Hickinbottom J (as he then was) in *Wilkes v DePuy International Ltd* [2016] EWHC 3096 (QB) [2017] 3 All ER 589 (“*Wilkes*”) at [71] that the test of what persons generally are “entitled to expect” does not benefit from being re-described in that way. The dangers of doing so are self-evident when the reformulation omits the essential word “entitled” and adopts a phrase which, in this jurisdiction, is used as a term of art in a very different context.
96. It is important to bear in mind that the test is not that of an absolute level of safety, nor is there an absolute liability for harm caused by a harmful characteristic. The Act does not impose a warranty of performance on a producer: *Pollard v Tesco Stores Ltd* [2006] EWCA Civ 393 at [17]. All hip prostheses will eventually wear out and fail, if the patient survives long enough, and some will fail within 10 years: the natural propensity of a hip implant to fail therefore cannot be a “defect,” any more than the inevitable wear and tear that causes minute particles of debris to enter the patient’s body. Otherwise all hip implants would be “defective”, irrespective of the materials used in the articulation.
97. One might think it self-evident that if a product fails to meet the objective safety standard set out in Section 3 or Article 6, the defect is whatever it is about that product (its state, or condition, or the risks to health and safety or property that it poses) that leads the Court to conclude that it fails to meet the safety standard. In their final submissions the claimants contended that the appropriate approach to be taken by the Court was first to consider whether the product was defective, and only after answering that question in the affirmative, to proceed to identify the “defect” for the purposes of applying the causation test in Article 4 and any development risk defence. They suggested that it was not necessary to describe the defect until the subsequent part of the liability analysis is reached.
98. I cannot accept that approach, which disregards the burden on the injured party of proving the defect as well as the causal relationship between defect and damage. As Hickinbottom J observed in *Wilkes*, proof of a causal connection between defect and damage cannot rationally or even conceptually be attempted without ascertainment of whether there is a defect, and if so, what that defect might be. A producer is only liable under Article 1 for damage caused by a defect in his product. If there is no defect, the claim must fail. Section 2 of the Act sets out what the claimant must prove, consistently with Articles 1 and 4 of the Directive. Section 3 of the Act sets out the yardstick of safety by which a defect is established. One cannot divorce the defect from the concept of defectiveness in the manner suggested; they are two sides of the same coin.
99. In order to prove the defect, a claimant must establish what it is about the state or behaviour of the product or the risks that it posed that led it to fall below the level of safety that persons generally were entitled to expect at the time the product entered the market, although he need not prove the precise mechanism by which it came to

fall below that yardstick. The fact that a product fails following normal use and in circumstances in which a standard product would not have failed may suffice for the Court to draw the inference that it is defective, see e.g. *Ide v ATB Sales Ltd and Another* [2008] EWCA Civ 424. Thus, for example, if an electrical appliance bursts into flames if it is left plugged in, or a fridge explodes, it plainly does not meet the standard of safety that persons generally are entitled to expect, and it is unnecessary for the claimant to establish what caused it to catch fire or explode.

100. However, there may be circumstances in which a greater degree of specificity about a feature or characteristic that is said to make the product unsafe is required in order to prove the requisite lack of safety. That may be the case if the injury or damage complained of could have arisen even if the product met the objective standard of safety set out in s.3 of the Act, for example, in consequence of the manifestation of a known risk that could arise in normal use. Thus the claimant may have to establish that the failure of a product or a component in it was not due to ordinary wear and tear, but to something abnormal that caused it to fail when it should not have done; or that something must have happened to elevate the inherent risk to a level that was higher than the public was entitled to expect. *Wilkes* was an example of such a case, though on the facts the claim failed.

### **2.3 THE CLAIMANTS' PRIMARY CASE ON THE IDENTITY OF THE DEFECT**

101. The claimants' pleaded primary case is that the "tendency or propensity" of the Pinnacle MoM prosthesis to result in identified harm (ARMD leading to early revision) constituted a "defect" when considered by reference to all the relevant circumstances. Mr Oppenheim QC submitted that what makes a product defective is its inherent *potential for damage*; or a harmful characteristic. The more detailed features of the case, including the degree of harm suffered, (in this case, the allegedly high incidence of early revision) amount to the relevant circumstances against which the Court decides whether the product was in fact defective, but do not themselves constitute the defect. He submitted that there is nothing in the Act or the Directive to preclude a defect from being characterised as the product's potential for damage, that it accords with the objectives of the Directive to do so, and that the CJEU had adopted such a characterisation in *Boston Scientific Medizintechnik GmbH v AOK Sachsen-Anhalt-Die Gesundheitskasse* (Case C/503/13, 504/13) [2015] 3 CMLR 173 ("*Boston Scientific*").
102. The claimants' primary case involves the court taking the following approach:
- i) It must identify a harmful characteristic (or potential for damage) in a product;
  - ii) It must consider all the circumstances (though the claimants say the relevant circumstances are circumscribed);
  - iii) It must decide if the circumstances render the product defective (i.e. it falls below the level of safety the public is entitled to expect);
  - iv) If the product is defective, the harmful characteristic *becomes the defect*.
  - v) To establish causation, one asks if, on the balance of probabilities, harm would have occurred if the product had not been defective.

103. Mr Oppenheim submitted that the approach of Burton J in *A v NBA* (notwithstanding that it was expressly confined to what Burton J described as “non-standard” products) was consistent with the claimants’ primary approach to defect, in line with the approach advocated in *Boston Scientific* and *Sanofi Pasteur*, and in accordance with the purposive construction of the Act and Directive. Insofar as Hickinbottom J. decided differently in *Wilkes*, his observations were *obiter*; and in any event this Court should follow the earlier decision, though neither domestic authority is binding on it.
104. Mr Antelme QC, on behalf of DePuy, submitted that the claimants’ primary case on defect was fundamentally misconceived. He adopted the criticisms of the approach in *A v NBA* articulated in *Wilkes*. He rightly accepted that there was nothing objectionable, as such, in defining a “defect” as something about a product that creates a real risk of injury or as its potential for causing harm. DePuy’s objection was to the characterisation of a potentially harmful characteristic of the product that could arise during normal use and without there being anything wrong with the product as a “defect”, leaving out of account the abnormality that caused the alleged lack of safety complained of.
105. Mr Antelme submitted that the claimants’ primary case is based on a misinterpretation of the CJEU decision in *Boston Scientific*, which does not support it; in any event that decision was fact-specific to the products in question and not seeking to lay down any general propositions of law. There is nothing in the Directive or its objectives that requires a normal characteristic or something giving rise to a normal risk of harm to be treated as a defect, and that to do so would undermine the need to prove causation and make significant unwarranted inroads into the development risk defence. He submitted that the reasoning in *Wilkes* was correct and should be adopted.
106. As Mr Antelme pointed out, the Act and the Directive say nothing about any “harmful characteristic” and no such wording exists within the legislative framework; the Court should avoid the temptation to apply a different test from that which is plainly set out in the Act. Attempting to paraphrase or re-interpret the plain language of the Act (and the Directive) is a course fraught with danger. The approach advocated by the claimants is circular, because it requires the identification of a “harmful characteristic”, which in and of itself suggests a defect, and then asking whether in all the circumstances that harmful characteristic does constitute a defect. Moreover, making the first step to “identify the harmful characteristic which caused the injury”, as Burton J did in *A v NBA*, would be to identify the primacy of causation before any investigation of defect, which is contrary to the Act and the Directive. It involves reasoning backwards from the harm (or incidence of harm) to find a defect in a normal characteristic of the product, even though that harm may have occurred without the product being defective. It ignores entirely the central question of the expectation of safety that persons generally were entitled to have of the product.
107. In fact, once the defect is identified as the propensity of the product to cause a certain type of damage, then causation becomes virtually self-evident, as the claimant will have identified at the outset the characteristic of the product and the causal relationship between that product and the harm, and all the individual claimant would need to do was prove that he suffered harm of that type (in this case, that he or she underwent revision surgery caused in whole or in part by ARMD). Thus, on the claimants’ approach, once the defect is identified as the “harmful characteristic” the

answer to question (v) in the claimants' analysis will inevitably be "yes". Moreover, in such a case the producer would never be able to raise a development risk defence.

108. Like Hickinbottom J, I have substantial difficulties with each stage of Burton J's analysis. The approach adopted in *A v NBA* and advocated by the claimants in this case is self-evidently circular. I agree with Hickinbottom J's analysis at paragraphs [60] to [65] of *Wilkes*, and I prefer the approach that he took, which in my judgment is the correct one. I consider the claimants' primary approach to be directly contrary to the spirit and objectives of the Directive and the Act, which require the Court to move away from the concept of a "defect" being some flaw or failing in the product, and to concentrate instead on whether in all the circumstances the product meets the requisite objectively assessed safety standard.
109. I could find nothing in the objectives of the Directive (let alone the travaux) that supported the claimants' primary case, and the points made by Mr Antelme about causation and the development risk defence, as well as the approach of the CJEU in *Sanofi Pasteur*, pointed strongly against it. The imposition of no-fault liability on the producer of a defective product does not require the Court to define "defect" in any terms other than by reference to the test of safety set out in Article 6 of the Directive and reflected in s.3 of the Act.
110. As Hickinbottom J pointed out in *Wilkes*, safety is inherently and necessarily a relative concept, because no product, and particularly a medicinal product, if effective, can be absolutely safe. Even such commonly prescribed medicines as penicillin or aspirin can cause a hypersensitive response in certain patients which, in an extreme case, can prove fatal. The public is not entitled to expect that a product which is known to have an inherently harmful or potentially harmful characteristic will not cause that harm, especially if (as in the present case) the product cannot be used for its intended purpose without incurring the risk of that harm materialising.
111. The European Commission has accepted that there will be some products that would, by their very nature, carry some known risk of harm or damage (or, in other words, have a "harmful characteristic") but which cannot be regarded as defective for that reason alone. Viscount Davignon, answering a question from a MEP, stated:
- "The Commission agreed with the Honourable Member that nobody can expect from a product a degree of safety from risks which are, because of its particular nature, inherent in that product and generally known, e.g. the risk of damage to health caused by alcoholic beverages. Such a product is not defective within the meaning of .. the .. Directive".*
- That exchange was referred to in *A v NBA* at [31]. Burton J commented: *"this does not of course amount to an exemption for such a product from the article but simply an explanation of how the article operates."*
112. However, if the incidence of that harm, either in nature or degree, is abnormal, then the product *may* be regarded as falling below the standard of safety that persons generally are entitled to expect. If that is the case, the defect is not the inherently harmful characteristic which is part of the normal behaviour of the product, for so to characterise it would be to make all products of that type "defective," as they all bear that characteristic. That is the fundamental flaw in the claimants' original formulation.

The defect is the abnormal potential for harm, i.e. whatever it is about the condition or character of the product that elevates the underlying risk beyond the level of safety that the public is entitled to expect. That approach is not a blurring of the distinction between relevant circumstances and defect as Mr Oppenheim contended; it is identifying what it is about the condition or state of the product that makes it unsafe by the objective yardstick set out in s.3 of the Act.

113. Mr Antelme gave the following example of how the claimants' primary case might produce irrational results. Suppose that drug X normally has a 1 per cent risk of causing a stroke, but when taken in combination with drug Y it would cause an 80 per cent risk of causing a stroke. The effect of that combination is something that on the most advanced state of scientific knowledge at the time drug X was put on the market was unknown and could not have been discovered by the producer. If the "defect" is characterised as the propensity of drug X to cause a stroke, the producer would have no development risk defence, because that propensity was known. However, if drug X had no underlying risk of causing a stroke, but when taken in combination with drug Y there was an 80 per cent risk of causing one, there would be a potential development risk defence. Yet the risk to health and safety of the consumer which leads to the finding that the product is defective is precisely the same in each scenario.
114. As Mr Antelme put it, this creates a wholly irrational distinction between an unacceptable *increase* in risk and an unacceptable *creation* of risk. In both scenarios, it is the combination of the two drugs that brings about the unacceptable risk to health; it is the inability of drug X to be taken safely with drug Y, rather than the inherent propensity of drug X to cause a stroke that creates that risk and makes it defective, and the error in the claimants' analysis lies in treating the inherent propensity to cause a stroke as the defect. Once that error is eliminated, the producer would be able to run a development risk defence in both scenarios, because the focus would be on whether the abnormal safety risk was one that could have been ascertained on the state of scientific knowledge at the time that drug X was put on the market.
115. Mr Oppenheim's riposte was that in Mr Antelme's example, the harmful characteristic or potential for damage that makes the drug defective would not be the potential of drug X to cause a stroke, but the potential to cause a stroke when taken in combination with the other drug. I agree, but that proves DePuy's point. The combination giving rise to the elevated risk is not part of the circumstances to be ignored in characterising the defect. If there is something that increases an inherent risk to the point that an otherwise safe product falls below the standard of safety that persons generally are entitled to expect, the defect is not the inherent risk, but whatever gives rise to the abnormal, or increased, risk. One cannot ignore the abnormality or the elevated risk when characterising the defect.
116. It was known at the time that the Pinnacle Ultamet liner was launched in 2002 that all hip prostheses will shed debris in the course of their normal operation, and that all MoM articulations will produce some metal wear debris. It was also known that an adverse reaction to particulate debris, in general terms, could cause a hip prosthesis to fail and require revision. The public has no entitlement to expect that a MoM hip will not produce metal wear debris as a concomitant of normal use, any more than it would be entitled to expect a MoP hip not to produce polyethylene and/or metal debris, even though both types of debris may cause an adverse immunological reaction leading to

damage to tissue or bone, and that damage may be worse or arise more quickly in some patients than others.

117. For a small minority of patients with a MoM hip prosthesis, their immune systems will trigger a reaction to metal debris which can cause soft tissue damage necessitating a revision. They cannot be identified in advance. The development of ARMD is therefore one of the normal risks inherent in the use of the product. The natural propensity of a MoM hip prosthesis to shed metal debris that can cause soft tissue damage is impossible to eliminate, in just the same way as the propensity of a MoP hip to shed plastic debris that can cause bone resorption and osteolysis. If the propensity to cause ARMD or the inherent risk of causing ARMD were treated as a defect, then all MoM prostheses would be defective, irrespective of design or the amount of wear particles they produce, and no manufacturer would be able to avail himself of the development risk defence because that risk was a known one.
118. Quite understandably Mr Oppenheim balked at the suggestion that the claimants' primary case was that all MoM hip prostheses were defective because they have the potential to cause ARMD leading to revision surgery. At one point he submitted that they were not contending that the "potential for damage" was in itself the defect, but that such a "potential for damage" must be "abnormal" (even though the word "abnormal" appears nowhere in the claimants' pleaded primary case) and that the test of whether the inherent potential or risk of harm is "abnormal" is that set out in Article 6 of the Directive, namely, whether in all the legally relevant circumstances, the product provides the safety that persons generally were entitled to expect. He submitted that the incidence of risk of early failure for ARMD is what makes the inherent potential for causing ARMD "abnormal", but in characterising the defect one ignores the manifestation of (or calculation of) the risk, because it is part of the relevant circumstances. The defect thus becomes the inherent harmful characteristic or the potential for causing damage.
119. Unfortunately, that approach does not cure the problem that the logical consequence of Mr Oppenheim's argument is that all MoM hip prostheses are defective. If one ignores the abnormality in characterising the defect, and treats it merely as a "circumstance" that makes the product unsafe, then in theory, anyone with a MoM hip prosthesis who develops ARMD would be able to recover damages under the Act, even though at the onset of the analysis it must be assumed that they could develop ARMD without the prosthesis being defective. That is, in effect, the same circular argument that Burton J accepted in *A v NBA*, and which I regard as fundamentally flawed, for the same reasons as Hickinbottom J in *Wilkes*. The claimants are seeking to use the actual or predicted incidence of manifestation of a known inherent risk, under normal circumstances of use, to characterise something that is normal as a "defect". That approach is both misconceived and contrary to the spirit and intention of the Act and the Directive.
120. The claimants contended that the CJEU has conceptualised the "defect" in a product as its "potential for damage" and that when its potential for, or risk of causing damage becomes "abnormal" in the sense that it crosses the threshold of what persons generally are entitled to expect, the product is to be regarded as defective. In support of that proposition they relied on the decision in *Boston Scientific*, and in particular on observations made in the Opinion of the Advocate General in that case. That argument is based on a misinterpretation of *Boston Scientific*. It is not a case in which

the CJEU decided that the concept of “defect” under the Directive required interpretation, let alone attempted to define it in a manner that supports the claimants’ primary case. The “defect” had already been identified, and even accepted by the defendants, and the CJEU was not being asked to express any views about it.

121. The two cases referred to the CJEU by the German courts concerned cardiological products, defibrillators and pacemakers, each of which belonged to a product group in which some, but not all, the devices suffered from a fault which might cause the device to fail. In the case of the pacemakers, the likelihood of failure was 17-20 times greater than normal, and the manufacturer had recommended that they should be surgically removed from patients. For both products, if the problem manifested itself, the consequences for the patient could be fatal, and it was impossible to ascertain whether the device actually contained the identified fault whilst it was still implanted in the body. Therefore, on any view, some of the products within each group, being potentially lethal, plainly did not meet the objective standard of safety set out in Article 6 of the Directive.
122. The main issue for the CJEU to determine was whether it was sufficient to found liability for a patient who had been implanted with a device from within the product group, and who had undergone a prophylactic operation to remove the device, to prove that pacemakers or defibrillators within that product group had a significantly increased risk of failure; or whether it was necessary for them to prove, as the defendants contended, that the specific device was one of the devices within the product group that suffered from the identified fault.
123. The Advocate General, in his Opinion [at AG 30] expressed the view that “*the concept of safety which a person is entitled to expect, which is relatively imprecise and of indeterminate content ... must be understood to refer to a product that poses risks jeopardising the safety of its user and having an abnormal, unreasonable character exceeding the normal risks inherent in its use. Accordingly, the lack of safety does not stem from the danger that may be posed by the use of the product, as a product may be dangerous even without having a safety defect, but from the abnormal potential for damage that the product could cause to the person or to the property of its user. In other words, the defect for the purposes of art 6(1) of Directive 85/374 is a risk of damage to such a degree of seriousness that it affects the public’s legitimate expectations insofar as concerns safety*” (emphasis added).
124. He went on to ask, rhetorically, at [AG 33]:  
  
“*how could the public not have legitimate grounds for questioning the safety of a product that has exactly the same characteristics as other products which have been proven to have a significantly higher than normal risk of failure or in which failures have already occurred in significant numbers? From the point of view of users, it goes without saying that if a product’s design and manufacture are identical to those of other products, that product is treated in the same way as the others as regards their risk of failure.*”  
  
He concluded by observing, at [AG 55]:  
“*the fact that a certain product belongs to a defective product group suggests that it has potential for failure itself, which is at odds with what a person is entitled to expect as regards patient safety.*”



125. The CJEU found that in the light of the function of the defibrillators and pacemakers, the safety requirements which the public was entitled to expect were particularly high, and (endorsing what the Advocate General said in paragraph 30 of his Opinion) for products such as those at issue, the potential lack of safety arose from their abnormal potential for damage (i.e. if they failed, someone could die). Accordingly, it held at [41]:

*“where it is found that such products belonging to the same group or forming part of the same production series have a potential defect, it is possible to classify as defective all the products in that group or series, without there being any need to show that the product in question is defective”* (emphasis added).

126. That conclusion is hardly surprising. Once it was established that there was a group or class of medical implants, any one of which had a significantly elevated risk of failure which, if it occurred, could cause a patient to die, the group or class as a whole could not be said to have achieved the level of safety that the general public would expect, and therefore the same would be true of any product within that group. That is all that *Boston Scientific* decided. The defect in that case was not the prospect that a defibrillator or pacemaker might fail, but the *significantly increased risk of failure* from which the specific product groups suffered. That increased risk was sufficient to constitute a defect in those circumstances, for all products within the group, particularly because of the serious consequences if the risk materialised.
127. *Boston Scientific* was not a case about the natural risks inherent in the use of the product under normal conditions for the purposes intended. The potential lack of safety which gave rise to the liability of the producer arose from “*the abnormal potential for damage which those products might cause to the person concerned*” not the “*normal potential for damage seen within all the [relevant] circumstances.*” There is nothing in *Boston Scientific* to support the argument that the normal risks inherent in the use of a product can constitute a “defect”. Indeed, the observations of the Advocate General in the passage of his Opinion expressly adopted by the CJEU are to the opposite effect.
128. The claimants also relied on *Sanofi Pasteur*, which was a case about the lawfulness of the procedural requirements for proving causation in France. Though this was unnecessary to its decision, in paragraph 41, the CJEU considered how a person who had been administered a vaccine and subsequently contracted a disease might be able to discharge the burden of proving that the disease was due to a defect in the vaccine. The focus in that paragraph is on establishing the necessary causal link between defect and damage, not the characterisation of “defect”. The CJEU observed that the burden of proof might be discharged where the Court considers that “*the administering of the vaccine is the most plausible explanation for the occurrence of the disease and....that the vaccine therefore does not offer the safety that one is entitled to expect, taking all circumstances into account... because it causes abnormal and particularly serious damage to the patient who, in the light of the nature and function of the product, is entitled to expect a particularly high level of safety.*” There is then a reference to *Boston Scientific*.
129. That passage does not advance or endorse the claimants’ primary case. The CJEU was simply referring to a straightforward application of the test in Article 6 of the Directive to a pharmaceutical product, in that case, a vaccine, by pointing out that if a

legitimate inference can be drawn from the evidence that the vaccine was the most likely cause of a patient suffering abnormal and particularly serious damage, it may well be held to fall below the particularly high standard of safety that the public is entitled to expect of a vaccine. What made the product defective in that hypothetical example was its unknown potential to cause *abnormal* damage, in the form of a particular disease, which was self-evidently not among the safety risks that the public generally might be expected to accept when being vaccinated. There is nothing in *Sanofi Pasteur* that supports the suggestion that when identifying the defect, one can or should take out the abnormality.

130. The final authority relied upon by Mr Oppenheim was a domestic one: *Baker v KTM Sportmotorcycle UK Ltd and Another* [2017] EWCA Civ 378. The case concerned a motorcycle accident which was caused by the front brake of the claimant's relatively new motorbike seizing without warning. The reason for the seizure was found to be galvanic corrosion, which the judge at trial decided must have been due to a fault in the design or manufacturing process. The defendants challenged that finding on the basis that there was insufficient evidence to support it. At [36] Hamblen LJ stated that "the defect found was a susceptibility for galvanic corrosion to develop in the front brake system when it should not have done, i.e. after limited and normal use and notwithstanding proper servicing, cleaning and maintenance" and again at [39], "the brakes were defective in that they allowed galvanic corrosion to develop following normal use in circumstances where standard non-defective brakes would not have done" (emphasis added).
131. Thus, *Baker* was not a case about the normal incidence of galvanic corrosion in brakes, and the defect was not the normal susceptibility of motorcycle brakes to corrode, it was the *abnormal* susceptibility of the brakes on the claimant's motorcycle to corrode. Mr Oppenheim submitted that the fact that the corrosion occurred when it should not have done was part of the relevant circumstances, and the defect was correctly characterised as the potential for galvanic corrosion. That submission ignores the abnormality that was the essential aspect of their condition that made those brakes unsafe – the elevated risk of corrosion leading to early failure. The case does not support the proposition that the normal propensity of a product to cause harm, or the inherent risk of causing a particular type of harm in normal use, can be legitimately characterised as a "defect".
132. I also reject Mr Oppenheim's contention that the approach advocated by DePuy would extend the bounds of the development risk defence and transfer risk in a manner inimical to the Directive. On the contrary, the claimants' primary case would unduly narrow the boundaries of the development risk defence and transfer a risk to the producer that it was never intended he should bear, as well as making the need to prove causation virtually redundant.
133. For all the above reasons, the claimants' primary case on defect is misconceived and there is nothing in EU or domestic case law or in the Directive or the Act, let alone the travaux, to support it. The public is not entitled and was not entitled at the material time to expect that a MoM hip, or the Pinnacle Ultamet prosthesis in particular, would not shed metal debris, even though this could cause ARMD in some of those in whom it was implanted. The vast majority of patients who have received a Pinnacle MoM implant have well-functioning hips. The alleged incidence (or predicted incidence) of ARMD in a small minority of patients, even if established, cannot and does not turn a

known risk which may eventuate in normal use, into a “defect” as defined in s.3 of the Act. The propensity of a MoM hip to bring about ARMD is not a “defect”.

134. Although at one time the claimants appeared to suggest otherwise, there is no feature of a Pinnacle Ultamet prosthesis that makes it any more likely than any other MoM articulation to shed metal debris, let alone increase the risk of developing ARMD in patients receiving such an implant. On the contrary, the engineering evidence indicated that insofar as this was capable of being predicted at all, given the numerous complex factors that have a bearing on wear rates, then, by reason of its design features, the average level of wear debris likely to be generated from a Pinnacle 36mm head MoM bearing in normal use was lower than the average to be expected from large head MoM hips generally. The statistics on which the claimants rely are consistent with that; the cumulative risks of revision for the Pinnacle Ultamet prosthesis were lower than those for all large head MoM implants.
135. I propose to say nothing further in this judgment about the claimants’ primary case on “defect” because, for the reasons I have given, it is untenable.

#### **2.4 THE CLAIMANTS’ ALTERNATIVE CASE ON THE IDENTITY OF THE DEFECT**

136. The claimants’ alternative case is that the Pinnacle Ultamet prosthesis had “*an abnormal potential for damage, compared with existing established non-MoM total hip replacement prostheses and/or that a large head MoM articulation had an abnormal potential for damage compared to alternative bearing surfaces within the Pinnacle modular system*” [see the document entitled “Claimants’ Clarification of Case on Defect” dated 18 December 2015, served pursuant to an Order made at the CMC on 12 November 2015]. Mr Antelme raised no objection to this formulation as a matter of law, though he contended that the claimants were unable to prove the alleged defect.
137. In this case what the claimants are really complaining about in terms of safety is not that a large head MoM articulation can cause ARMD, but that because of the incidence of ARMD that they say materialised or can be reliably predicted, Pinnacle Ultamet prostheses had a materially higher risk of early failure than comparator non-MoM prostheses. If it is established on the evidence that the Pinnacle Ultamet prostheses did indeed have a materially greater risk of early failure and revision, when compared with an appropriate comparator prosthesis or prostheses that was or were on the market at the relevant time, the “defect” is not the natural propensity of the product to cause ARMD, but its abnormal risk of failure for ARMD leading to early revision when measured against that comparator, which is the claimants’ alternative case. The claimants’ alternative formulation is consistent with the approach in *Wilkes* and *Boston Scientific*.
138. I will consider the merits of the alternative case and the circumstances on which the claimants and DePuy rely in due course, but before doing so, it is necessary to resolve the dispute between the parties about whether there are any circumstances, other than fault, that the court is obliged to leave out of account when evaluating the safety of a product.

## **2.5 LEGALLY RELEVANT CIRCUMSTANCES**

139. The Act and the Directive require the Court to have regard to “all the circumstances” and in particular to the non-exhaustive list of circumstances set out in Article 6(1), reflected, and to some extent amplified, under s.3 of the Act. The claimants submitted, and I accept, that the specified circumstances, whilst not exhaustive, are to be regarded as the most significant. I also accept, as Hickinbottom J did in *Wilkes*, that the circumstances to be taken into account in a given case must be factually and legally relevant to the evaluation of safety.
140. Of course, as Mr Antelme accepted, the question whether the producer was at fault is legally irrelevant, since the Directive has introduced a system of no-fault liability; a producer may even be held liable for a defect that he took all possible precautions to prevent, subject to the development risk defence. Fault, or absence of fault, is also *factually* irrelevant, because it plainly has no bearing on the condition or state of the product that makes it fall below the standard of safety that the public is entitled to expect.
141. In *A v NBA*, at [68] Burton J expressed the view that the following circumstances were not legally relevant, at least when evaluating whether a non-standard product is defective: the avoidability or impossibility of taking precautionary measures; the impracticality, difficulty or cost of taking such measures; and the benefit to society or utility of the product (except in the context of whether, with full information and proper knowledge, the public does and ought to accept the risk). In the present case, the claimants contended that Burton J was not only right to exclude such matters from consideration when determining whether non-standard products are defective, but that these factors were also irrelevant considerations when evaluating whether standard products such as the Pinnacle MoM hip prostheses met the entitled expectation of safety. They submitted that there are certain circumstances that would never be relevant legally because they fall outwith the purpose of the Directive. Thus avoidability, cost, and any benefits other than those specifically relating to safety must *always* be excluded from consideration. So too must the compliance of the product with any regulatory requirements or standards – even if those requirements have a direct bearing on the aspects of the product’s safety of which complaint is made.
142. The claimants submitted that any consideration of these matters would shift the balance away from the consumer in favour of the producer and would create the danger of introducing negligence-type concepts into a strict liability regime. They also submitted that this would run contrary to the objective of providing a more simplified and straightforward means of redress for the consumer, although, as I have already explained, I do not accept that this was an objective of the Directive beyond the simplification necessarily brought about by the introduction of a no-fault liability regime.
143. In general, I prefer Hickinbottom J’s approach in *Wilkes* to that of Burton J in *A v NBA*. I agree with Hickinbottom J’s observations in *Wilkes* at [77] to [79], especially that:  
“the court must maintain a flexible approach to the assessment of the appropriate level of safety, including which circumstances are relevant and the weight to be given to each, those factors being quintessentially dependent upon the particular facts of any case.”

**AVOIDABILITY, RISK-BENEFIT AND COST**

144. Like Hickinbottom J in *Wilkes*, I consider that these factors can usefully be discussed together. The use of the label “risk-benefit” carries with it a danger of confusion. Neither DePuy nor any of the other non-party producers took the position that the Court should make a list of the risks and a list of the benefits and weigh one against the other. They were not advocating the adoption of a US-type system, which the framers of the Directive had specifically rejected. The issue in this litigation, as it was in *Wilkes*, is whether it is legally permissible for the Court, when evaluating whether the safety of a product was that which the public generally is entitled to expect, to take into account any benefits it brings (other than benefits directly related to safety) when such benefits would be factually relevant to that evaluation, or whether the Court is constrained simply to look at the risks the product poses.
145. I appreciate that in many cases the question whether the risk of harm could have been avoided, or reduced, or the benefits that the product might confer, for example, in terms of its utility, would be irrelevant considerations (or at best, would carry very little weight) in the overall evaluation of whether it meets the requisite safety standard. I accept, as Burton J accepted, that it is not a defence under the Directive or the Act for the producer to demonstrate that the risk was unavoidable, and that there may be situations in which a producer will be held liable for defects that he could do nothing to prevent.
146. However, that will not necessarily always be the case. Whilst the ability or inability of the producer to eliminate the safety risk posed by a product is unlikely to have any relevance to an assessment of how safe the product is, it is possible to envisage circumstances in which it might legitimately form part of a holistic evaluation of whether that level of safety falls below the threshold set in s.3 of the Act – i.e. the level that the public generally was *entitled* to expect. Inquiring whether there was fault or absence of fault on the part of the producer is not the same thing as considering whether, as a matter of fact, the risk of injury could have been eliminated or reduced, and whether that has any impact on what the public generally was entitled to expect in terms of its safety.
147. I sympathise with Burton J’s view that in a case such as *A v NBA* the question whether the harm could have been avoided, and the impossibility of taking any or any further precautionary measures, or the prohibitive expense of doing so, would have no bearing on the key issue of whether the product met the objective safety standard. It is nothing to the point how the contaminant came to be in the product, if the presence of the contaminant (with or without the absence of warnings) would be sufficient to make the product fall below the yardstick of safety set by s.3. Once it falls below that yardstick, then it is defective, and then it will be for the producer to establish any defence, such as the development risk defence. Likewise, if a product which a patient has no choice but to use has been contaminated by something inherently harmful that one would not expect to find in it, obviously the way in which an *uncontaminated* product would behave and the benefits that a blood transfusion could confer is unlikely to be informative of the safety of the contaminated product (save possibly as a comparator).
148. Indeed, I can envisage many cases in which the question whether the risk could have been avoided, or a consideration of benefits unrelated to safety, would play no role in

the overall evaluation of the safety even of a product which is made to specification and subjected to normal use. For example, if a drug manufactured to specification manifests an unpredictable and unavoidable side-effect which causes brain damage, it would almost certainly fall below the level of safety that the public was entitled to expect, and the inability of the manufacturer to have prevented that side-effect, or the benefits it might have conferred on patients if it did not cause them brain damage, would plainly have no impact on the evaluation of the entitled expectation of safety.

149. In *Sanofi Pasteur*, addressing that type of scenario, Advocate General Bobek observed at [87] (albeit in a passage that the CJEU did not adopt) that the objective standard of safety in Article 6 “*essentially refers to baseline expectations of the product under normal conditions of use. It does not mean that where the product is used normally and causes serious harm in an individual case, that a conclusion of defectiveness necessarily requires a balancing of the costs and benefits of the product.*” The word “necessarily” is important; the Advocate General was not expressing a view that costs or benefits are always irrelevant considerations.
150. It is important to put those observations into context. In that passage, the Advocate General was considering (and rejecting) the submission by one of the respondents that the elements establishing a causal link between product and damage in an individual case cannot alone suffice to establish defectiveness, but that a broader assessment of the costs/benefits of the product is *always* required, going beyond the specific case. The respondent was contending that a specific product can be considered defective only if the product is more generally found to be unsafe: see [85] and [89]. That would require the claimant not only to prove that the product he used was defective but that there was a generic defect in all products of that type or batch. Addressing that argument, the Advocate General understandably concluded [at 88] that imposing such a requirement would amount to creating new conditions of liability. He rejected the suggestion, made in support of that contention, that cost and benefit must *always* be considered relevant, which is quite different from suggesting that they could *never* be considered relevant.
151. In my judgment, nothing can be drawn from the observations of the Advocate General (or the silence of the CJEU on this topic) in *Sanofi Pasteur*, or indeed from the understandable absence of any mention of risk-benefit in *Boston Scientific* to support the claimants’ submission that these factors are legally irrelevant in all cases or that Hickinbottom J’s observations in *Wilkes* about risk-benefit were misconceived. No-one could have seriously suggested in *Boston Scientific* that the benefits conferred on patients by a normally functioning pacemaker could have any relevance to the decision whether the 17- to 20-fold risk of cardiac arrest if the patient received one of the faulty pacemakers was sufficient to make those pacemakers defective, or that the limited residual activity that might be achieved by deactivation of the faulty switch in the defibrillators had any bearing on the acceptability of the risk of death if a faulty defibrillator was left in place untouched. Neither *Boston Scientific* nor *Sanofi* were cited to Hickinbottom J in *Wilkes*, but they do not compel the Court to draw the conclusion that his approach was wrong.
152. Depending on the product, I agree with Hickinbottom J that such factors *may* sometimes play a legitimate part in the evaluation of the degree of safety risk that the public would be expected to tolerate, because there may be cases in which one cannot fairly or sensibly evaluate the latter without looking at the former. I do not accept the

argument, at one point advocated by the claimants, that one never gets to consideration of benefit because one does not need to look further than the presentation and packaging of the product, and any warnings that have been disseminated. Of course, these are important circumstances to be considered, but they are not surrogates.

153. As DePuy submitted, where a product includes a feature which gives it a potential functional advantage, or eliminates a perceived deficiency in design, but by doing so necessarily introduces a risk, it is artificial to prevent the Court from considering that actual or potential benefit when making an assessment of whether the product is defective. Depending on the circumstances, including any warnings, the safety risk may be one that, objectively, the public would be expected to accept, bearing in mind the benefits that the product would confer. Much will depend on the nature and seriousness of the risk and the likelihood of its manifestation.
154. If Burton J's analysis had gone no further than stating that avoidability and benefit were irrelevant considerations in assessing whether the blood products with which he was concerned met the requisite standard of safety, it would have been unobjectionable. The problems with *A v NBA* lie in the passages that appear to lay down rigid rules of more general application pertaining to the circumstances that fall to be legally excluded from consideration when evaluating safety (at least in a non-standard product) because they are somehow deemed to undermine, or run counter to, the no-fault basis of liability. Burton J took the view that without excluding such matters, subject only to the development risk defence, the Directive would not only be toothless but pointless. I respectfully beg to differ. Dictating that certain circumstances can *never* be considered relevant in law or taken into consideration (however relevant they may be in fact), undermines the flexibility of the Act and the Directive, which caters for the fact that the suitable approach for one product may be inapposite for another.
155. Burton J expressly recognized that if his analysis were correct and followed in other cases, problems may arise in the consideration of what he termed a "standard" product, but he did not consider that such problems would be insurmountable if safety, use and the circumstances identified in Article 6 were kept at the forefront of consideration. It seems to me that an analysis which throws up problems, whether insurmountable or not, is best avoided, especially if it is unnecessary. There is no need to introduce any level of complexity or rigidity into the application of the statutory test. It is both difficult and unwise to attempt to categorise the circumstances that are or are not relevant to the assessment of safety of different kinds of products.
156. Moreover, if the Council had wished to exclude certain circumstances from consideration, even if they might otherwise have had a direct bearing on the assessment of what the public was entitled to expect in terms of safety, it would hardly have used the expression "all the circumstances". The Court must decide what is and is not relevant to the assessment, on a case by case basis, always by reference to the test of safety set out in s.3. Its hands should not be artificially tied.
157. The claimants submitted that the distinction between standard and non-standard products used by Burton J in *A v NBA* was simply a helpful analytical tool, and that as Burton J recognized, it may be easier to prove defectiveness if the product differs from the standard product in a material respect. They pointed out that in *Wilkes*

Hickinbottom J acknowledged that whether a product is within the producer's specification may be a relevant circumstance in relation to whether the level of safety is that which persons generally are entitled to expect. However, he also said that to raise the distinction between standard and non-standard products to a rigid categorisation is positively unhelpful and potentially dangerous. Therefore, whilst he did not disagree with the use of the distinction as an analytical tool in an appropriate case, he criticised its use to determine the approach to the circumstances which bear upon the definition of "defect." That is why he said, at [94], that the distinction was "*as a classification unnecessary and undesirable*" (emphasis added).

158. I, too, accept that the inquiry whether the product met its specification and was subjected to normal conditions of use when the harm or injury arose may be an appropriate one, since those matters plainly form part of the relevant circumstances that the Court is entitled to bear in mind when assessing safety. There is no reason, therefore, why it should not be adopted as a useful starting point in the analysis, provided that it is not treated as an obligatory approach in every case. It may be that a case involves products which cannot easily be classified as standard or non-standard (e.g. a prototype), or products falling into both categories.
159. If it is possible to classify the product as standard or non-standard, this classification may, but will not necessarily have a significant bearing on the circumstances that are relevant to the evaluation of its safety. Both Burton J and Hickinbottom J rightly acknowledged that it may be easier to prove that a product is defective if it is out of specification, but they also acknowledged that the mere fact that it is out of specification may not be enough to prove that it lacks the requisite degree of safety. Conversely, a product that meets its specification may still be unsafe when evaluated by the objective test in s.3 of the Act.
160. I agree with Hickinbottom J that what may be a useful analytical tool should not be elevated to a rigid categorisation which then dictates what circumstances are and are not to be considered when making the assessment of safety. Like him, and for essentially the same reasons as are set out in his judgment in *Wilkes*, I take the view that the Court is entitled to have regard to all circumstances which may have a bearing on the assessment of the safety of the product; that those circumstances may differ, depending on the product and the nature of the complaint about it; and that there is nothing in the Directive that compels the Court to disregard a circumstance which plainly has a bearing on the level of safety which the public was entitled to expect in a given case, either because the product is "non-standard" or because the regime is one of non-fault based liability. The Court simply needs to be vigilant not to let notions of negligence or other irrelevant considerations creep into that assessment.
161. As Hickinbottom J acknowledged in *Wilkes*, a pharmaceutical product that is highly beneficial to most patients but in a minority causes death or serious injury for a reason unascertained and unascertainable, may nevertheless be held to lack the appropriate level of safety. Yet in my judgment it would be wrong in principle to exclude from consideration of what level of safety the public is entitled to expect, the benefits that the product could confer, or to confine the relevant benefits to safety benefits, in cases in which those wider benefits might properly have a bearing on that assessment.
162. Mr Antelme gave the helpful example of a new chemotherapy drug which has proven advantages over all others on the market, but a rare and serious side effect. In



evaluating a claim made under the Act by a patient who has been harmed by suffering from that side effect, if the Court simply looked at the harm the drug could cause, and excluded its particular benefits, namely, the better prospect of successfully treating the cancer, it is difficult to see how it could properly evaluate the level of safety that the public is *entitled* to expect of that drug. Indeed, if the enquiry is so constrained, it might be difficult to avoid the conclusion that the product was defective, because whilst the use of the drug for chemotherapy would be taken into account, the Court would be excluding from consideration the features of the drug that make it more effective than other chemotherapy drugs which did not have that side effect.

163. The conclusion might therefore be drawn that the public is entitled to expect that they should not be exposed to the risk of a serious side effect which they would not suffer if they took a different drug which was (on the face of it) equally effective to combat the same cancer. When the benefits of the new drug are included in the assessment, the conclusion as to safety might still be the same, depending on all the other circumstances, including warnings, but it *could* be different – and, more importantly, the evaluation of safety and defectiveness would be made after taking all relevant information into account.
164. If the use to which the product can reasonably be expected to be put is a relevant consideration, as it undoubtedly is, then it cannot be objectionable for the Court to consider the benefits likely to arise from its contemplated use as part and parcel of the circumstances that have a bearing on the evaluation of the level of safety that the public generally is entitled to expect. In the example given, the additional benefit conferred by the new chemotherapy drug is plainly a relevant circumstance that would assist in the evaluation of its safety by reference to the test set out in s.3 of the Act.
165. Considering benefits and/or avoidability in an appropriate case is not importing a product's fitness for use into the analysis or introducing a concept from the US system that the legislators rejected, as the claimants submitted, but rather, taking a holistic approach to the objective evaluation of safety, as Field J did in *Bogle v McDonalds Restaurants* [2002] EWHC 490 (QB), [2002] 2 All ER (D) 436. That case concerned a claim by a group of claimants who had been burned by hot drinks served in various branches of McDonalds. One of the ways in which they put their case was that the hot drinks were defective products under the Act. In assessing the level of safety that the public was entitled to expect, Field J had regard to the benefit or utility of being able to remove the lid in order to consume a hot drink in the preferred manner, even though the ability to remove the lid created a greater risk of spillage and scalding than leaving it in place would have done. As Hickinbottom J put it in *Wilkes* at [88],
- “although the risk of scalding was avoidable in absolute terms, the cost of avoiding it in terms of utility was unacceptably high. Thus avoidability as seen in the broader context of the risk-benefit balance was, in substance, taken into account.”*
166. Of course, when considering avoidability, as Hickinbottom J pointed out, there is a danger of unduly focusing on the acts and omissions of the producer of the product rather than the product itself. The Court must bear in mind that the regime is not fault-based. That is why, even in cases in which it is relevant, avoidability should not be examined in a vacuum or given a prominence that it does not merit. I agree with Hickinbottom J that in respect of a medical product, such as a prosthesis, a detailed consideration of the discrete question of whether a particular risk is or is not avoidable

is unlikely to be fruitful. Nevertheless, I share his view that in an appropriate case and without inappropriately moving the focus of the exercise, the ease and extent to which a risk can be eliminated or mitigated *may* be a circumstance that bears upon the issue of the level of safety that the public generally is entitled to expect.

167. The Council considered that the economic risk of producing a defective product should fall on the person best equipped to take precautionary steps *or* to insure against it, namely, the producer. That is something the Court must not forget. Nevertheless, depending on the circumstances and the product, and particularly on the harm, it may be that in an appropriate case the Court could legitimately conclude that the public would not be entitled to expect the producer to achieve something in terms of safety that is scientifically impossible or prohibitively expensive or which it would be impossible to insure against. It all depends on the nature of the product and its intended use, and all the other relevant circumstances.

### **LEARNED INTERMEDIARIES**

168. So far as the presence of an intervening healthcare professional (or “learned intermediary”) is concerned, the claimants submitted that the focus of the Court should be on the entitled expectations of safety of the consumer (or body of consumers generally). They did not go so far as to suggest that Hickinbottom J was wrong to conclude in *Wilkes* at [108] that the presence of a learned intermediary was a relevant circumstance (indeed he said that the suggestion that information provided to a learned intermediary was irrelevant, was unarguable). However, they submitted that the fact of such an intermediary does not change the level of safety that the public generally is entitled to expect.
169. Again, I respectfully agree with Hickinbottom J’s conclusion that the existence of a learned intermediary and the information and warnings provided to that intermediary are plainly relevant circumstances, for the reasons that he gave; however, the weight to be given to the existence of a learned intermediary and the information, including warnings, passed on to such intermediary, in the evaluation of whether the product met the entitled expectation of safety will depend on the circumstances of the individual case. I agree with the claimants that in assessing safety, the focus must be on what the public generally are entitled to expect, not what clinicians are entitled to expect, but the latter may have a considerable bearing on the former. I also agree with the point made by Mr Antelme that where a product is not defective, the regime is not designed to classify it as defective because of some fault or failing on the part of the intermediary (for example, a failure to pass on warnings or obtain properly informed consent).

### **REGULATIONS AND STANDARDS**

170. The claimants accepted that the fact that a product is regulated and subject to a regulatory regime is a relevant circumstance but submitted that this fact only served to *heighten* the standard of safety which the public would be entitled to expect, and that the question whether the product actually met the standards set by the regulatory regime was a wholly irrelevant consideration. I was unimpressed by the artificial distinction they sought to draw between the fact of regulation, and compliance or non-compliance with the regulatory regime. It is not justified by anything that was said by the Advocate General, let alone the CJEU, in *Boston Scientific*. In that case it was

irrelevant to consider whether non-faulty defibrillators or pacemakers generally met certain safety standards, because the faulty products would not have done. In that context, it was hardly surprising that the Advocate General suggested that the relevance of the regulatory regime was that its existence supported or indicated a heightened expectation of safety for products of that type.

171. The claimants sought to rely on the fact that it is a defence under Article 7(d) for the producer to prove that the defect was due to compliance of the product with mandatory regulations issued by the public authorities, as an indication that this was the only context in which such compliance would be considered relevant. However, it cannot be inferred from the fact that this specific defence exists, and that Article 7(d) is the only place in the Directive that compliance with regulations is mentioned, that compliance or non-compliance with regulations or safety standards is to be treated as an irrelevant circumstance at the earlier stage of the analysis when considering whether the product is defective.
172. At most, the existence of the defence serves to underline the fact that a producer may be held to have produced a defective product *notwithstanding* compliance with regulations or standards. Indeed if, as the claimants accept, the existence of regulations or standards are material factors because they indicate that the product is of a nature that requires regulations or standards to be imposed on the producer, thereby generating a heightened expectation of safety in comparison to unregulated products, I can see no good reason for failing to ask the obvious question whether those regulations or standards have been met.
173. Nor am I persuaded that any inference can be drawn from the rejection of a suggestion, made before the Directive was finalised, that compliance with a standard should be treated as *prima facie* evidence of the absence of defect, referred to in passing by Burton J in *A v NBA* at [35]. No-one has suggested that compliance with standards or regulations affords a defence or creates any *prima facie* presumption in favour of the producer. The weight to be ascribed to these factors will depend on the facts and circumstances of the individual case.
174. The claimants contended that allowing compliance with a regulatory regime to be treated as a relevant circumstance would encourage an enquiry into matters such as whether the body in question that set the standards was in fact a regulator, whether it relied on self-certification, whether the producer provided that body with full information, and even whether the grant of approval was correct, all of which would draw the Court into fault-like enquiries that are contrary to the purpose of the Directive and undermine the intention that consumers should have a relatively swift and simple means of achieving redress. I cannot accept those submissions. The nature of the standard or regulation, the body setting it and the information on which it relied may have some bearing on the weight, if any, to be placed on compliance or non-compliance, but that does not involve the Court in second-guessing a regulatory decision, let alone considering questions of fault. The Court will not allow the parties to introduce factual disputes that will divert it from consideration of the safety of the product.
175. Of course, the standards set by a regulatory regime cannot be used as a substitute for the statutory test of safety, even when the regulatory regime expressly addresses safety. The level of safety that the public is entitled to expect may be lower than a

particular safety standard, as in *Pollard v Tesco Stores*, where the British safety standard for childproof caps on which the claimants sought to rely did not apply to dishwasher powder bottles, because there was no requirement at the time that such products should have a childproof cap. In other cases, it may be higher, for example if the product complied in all material respects with particular safety features required by the regulatory regime, but there was some additional feature that made it unsafe; or where the generic products complied with the regulatory regime but there was a faulty batch that would have failed the safety assessment, as in *Boston Scientific*. The weight to be placed on such compliance is a matter of fact and degree in the individual case, and it may be of no relevance at all. However, the distinction that the claimants seek to draw between the fact that a regulatory regime exists, and the fact that there has (or has not) been compliance with it is illogical and unjustified.

176. In *Wilkes*, unlike the present case, it was common ground that non-compliance with any appropriate mandatory standards will provide evidence of a defect, and that compliance with such standards, whilst not providing a complete defence, will provide evidence that, in respect of the matters to which those standards go, the level of safety required by the Act has been satisfied and the product is not defective. Hickinbottom J said, and I agree, that in an appropriate case compliance with such standards will have considerable weight, because they have been set at a level which the appropriate regulatory authority has determined is appropriate for safety purposes. However, the standards must have a relevance to the defect that is alleged; it would be no good establishing that a child's toy met all the safety standards in terms of toxicity, for example, if the complaint was that one of the components was a choking hazard.
177. *Wilkes* was a case in which it was contended that the stem of a hip prosthesis failed earlier than it should have done by reason of a fatigue fracture that developed in it. In that context, the compliance of the stem with all relevant mandatory standards pertaining to the load-bearing capabilities of the prosthesis, the fact that it was tested to a higher level than the fatigue failure standard, and that it satisfied all the regulatory requirements including those imposed to ensure that the product was acceptably safe, understandably had a considerable bearing on the assessment of safety.
178. In the present litigation, however, the complaint relates to something that is not directly addressed in any safety standard or regulation, nor could it be, because the state of science is such that it is impossible to set any product specifications which would have a direct impact on the incidence of ARMD. For that reason, compliance with regulatory requirements may not play as significant a part in the overall assessment of defectiveness as it did in a case such as *Wilkes*.

## **2.6 CAUSATION**

179. Neither the Act nor the Directive prescribe how causation is to be approached; the case of *Sanofi Pasteur* made it clear that this is a matter for the national courts, subject only to ensuring that the rules adopted do not undermine the scheme of no-fault liability and create a different balance between the competing interests from the one that is set by the Directive.
180. The claimants' alternative case depends on proof that the Pinnacle Ultamet prosthesis had an abnormal potential for damage in comparison with established hip prostheses and/or other bearing surfaces within the Pinnacle system. Their complaint is that the

Pinnacle Ultamet prosthesis carried a materially increased risk of early failure because of ARMD, requiring revision surgery. It is common ground that all hip prostheses have an underlying risk of failure, requiring revision, within 10 years. Therefore, each claimant must prove that the increased risk was what caused them to undergo the early revision, rather than the inherent risk of failure leading to early revision that would arise irrespective of that defect.

181. In a claim for personal injury, the conventional way of proving causation is by establishing that, but for the breach of duty, the relevant injury would not have occurred. There is no conceptual difficulty in transposing the conventional “but for” test from fault-based liability to a strict liability arising from statute: the claimant would simply need to prove that but for the identified defect he would not have suffered the injury. However, the application of the “but for” test is not entirely straightforward in a case where the defect is an abnormal or increased risk of an early failure which could occur even if there was no defect. On the face of it, in order to establish causation, each of the claimants would have to prove that they developed ARMD, that that was the reason for (or at least made a material contribution to) the early revision, and that they would not have had an early revision but for the fact that they developed ARMD. The difficulty potentially arises at the third stage.
182. Mr Antelme submitted that, in order to prove on the balance of probabilities that the revision in a given case was due to the *increased* risk of failure within 10 years, rather than the underlying risk of that failure occurring, it must be established that the risk of failure due to ARMD was at least double the risk of failure that arose in any event. In support of that proposition he relied on the case of *XYZ v Schering Health Care Ltd* [2002] EWHC 1420 (QB) and on subsequent authorities in which a test of “more than doubling the risk” had been adopted, including *Novartis Grimsby v Cookson* [2007] EWCA Civ 1261 and *Jones and others v Secretary of State for Energy and Climate Change* [2012] EWHC 2936.
183. In *Jones*, Swift J set out at 6.19-6.54 an exemplary synopsis of all the earlier cases in which that test was discussed, and either rejected or applied. However, that case was one in which the Court was faced with a far more complex exercise, involving two discrete potential causes of the personal injury. It had to ascertain the way in which the claimants could prove that their occupational exposure to a carcinogen had caused them to develop lung cancer, when they were already exposed to an underlying risk of developing lung cancer because they were smokers. Swift J was very careful to explain why she reached the conclusion that she could not use the conventional test to assess whether the occupational exposure to carcinogens made a material contribution to the development of cancer, and therefore that the “doubling of the risk” test was apposite, as she put it, “*in the circumstances of this litigation.*”
184. The claimants in this case took a different approach to causation, essentially advocating the application of the classic “but for” test. Mr Oppenheim submitted that it was important for the Court to be astute not to impose a causal requirement that made the exercise of a claimant’s rights extremely difficult in practice. He submitted that damage is established if there was a revision caused in whole or in part by ARMD, but that in order to prove causation, the Court has to ask what would have happened if the prosthesis implanted in the claimant did not have the identified abnormal potential for damage. Therefore, he submitted, the correct question for the Court to ask is whether a claimant would have suffered ARMD (and not whether the

claimant would have undergone a revision for an early failure) if he or she had had an implant that did not create the abnormal potential for damage of that type. He submitted that because the adverse reaction manifests itself in soft tissue damage, and for causation purposes the Court is concerned with the causal connection between that defect and the damage of which the claimant is complaining, the question is whether, if the claimant had a different type of prosthesis implanted, they would have been revised for ARMD.

185. By so characterising the test, the claimants were necessarily setting themselves up to succeed on causation merely by proving the defect and the damage, because although other types of prosthesis can shed metal particles in consequence of ordinary wear and tear, the incidence of ARMD even in MoP articulations is negligible, and in CoC articulations would not arise at all. However, that is not the reason why Mr Oppenheim's formulation is wrong in principle. The correct question for the Court to ask is whether, on the balance of probabilities, the claimant would have suffered the damage complained of, i.e. undergone an early revision, if the hip had not carried with it the increased risk of early failure which made it defective. That means the appropriate comparison is with the generic incidence of early failure, and not just failure for ARMD.
186. If it were to be established that the incidence of early failure of Pinnacle Ultamet hips by reason of ARMD was more than double the general incidence of early failure of comparator hip prostheses, then of course the claimant would discharge the burden of proving causation. However, I am not persuaded that this is the only way in which causation could be established, or that "more than double the risk" should be adopted as a bright-line test in a case such as this, which really does not throw up the same kind of difficulties that the Court was faced with in *Jones* and *Novartis*. In any event, it is unnecessary for me to reach a conclusion on this issue, as the problem of proving causation was unlikely to arise in the present litigation if the claimants proved that the Pinnacle Ultamet prosthesis carried with it a materially increased risk of early revision (on their case, three to five times greater) than a relevant comparator prosthesis.

## **CHAPTER 3: THE FACTUAL BACKGROUND**

### **3.1 A SHORT HISTORY OF MODERN PROSTHESES**

187. In the 1950s McKee and Watson-Farrar designed a stainless-steel prosthesis, building upon a model which was originally used by Philip Wiles in 1938 at the Middlesex Hospital in London. However, it suffered from early loosening, and they revised their design to utilise a cobalt-chromium alloy. Meanwhile, Sir John Charnley began to experiment with using a Teflon acetabular cup combined with a small head stainless steel monobloc femoral stem. This did not perform well, due to excessive wear, and the design was abandoned in 1961.
188. A major breakthrough came in the early 1960s when Sir John published a new design which he termed the "low friction arthroplasty," a cemented implant with a MoP articulation. The stem of the Charnley prosthesis consisted of a monobloc of cobalt chrome steel, with a small metal head, and the cup was a monobloc of conventional (ultra-high molecular weight) polyethylene. The Charnley hip was very successful and generally outperformed the various first-generation MoM designs that were being developed throughout the 1960s. The performance of the latter prostheses was very

mixed, with poor patient outcomes and relatively high failure rates in the first years of clinical use. The various reasons for the rate of failure of these older designs included flaws in the design and manufacturing processes, and problems with their fixation.

189. Up to the early to mid-1990s the Charnley hip was the prevailing articulation and it, or variants of it, are still in use today. However, its major drawback was the risk of failure due to aseptic loosening caused by an adverse reaction to particulate debris in the body that was largely produced by wear to the polyethylene liner (i.e. osteolysis).
190. During the early part of the 1970s first-generation alumina ceramic was introduced as a bearing surface as a harder-wearing alternative to conventional polyethylene. Rather like the first generation of MoM implants, despite some initial optimism, its early performance was very mixed. Problems arose from a relatively high fracture rate, as well as difficulties with aseptic loosening. Professor Pandit likened the fracture of a ceramic hip to a broken vase where it is difficult to see all the shards and remove them. Mr Whitwell explained that from a surgical perspective, the fracturing of a ceramic head or cup can cause very serious practical difficulties, not only because of the challenge for the surgeon in removing all the parts of the shattered component, but because of the risk that any parts left behind might interfere with the revision prosthesis. He said the lesson that it was necessary to do a very good clearance around the hip to get those particles out had unfortunately been learned from the incidence of early revision failures in patients who had experienced ceramic fractures. Both he and Professor Pandit also referred to a problem of squeaking, possibly caused by edge loading of the ceramic bearing, which for some patients could be intolerable (and lead to them requesting a revision).
191. By around the middle of the 1970s most first-generation MoM bearings were no longer used, as it was perceived that they were outperformed by the MoP designs. However, the search for a bearing surface that would prove to be more durable than conventional polyethylene continued. During the 1980s, second-generation alumina CoC bearings were developed. Whilst the incidence of fracture decreased, they continued to suffer from aseptic loosening. A second generation of MoM bearings were developed in the latter half of that decade. The Swiss company, Sulzer, produced a design which used 28 mm and 32 mm femoral heads made from a Co-Cr-Mo alloy. Meanwhile, the designers of MoP articulations experimented with the intentional structural cross-linking of conventional polyethylene through irradiation to produce a harder wearing material, which it was also hoped would be less likely to degrade over time.
192. The increased use of MoP designs, particularly the Charnley hip, led to an increased incidence of failure of such implants due to polyethylene wear osteolysis. Other modes of failure or complication of MoP prostheses existed, including dislocation, but by the mid-1990s aseptic loosening caused by osteolysis was widely recognised to be the predominant cause of failure, especially in younger or more active patients. The length of time before such failure occurred would depend on multiple factors, including how active the patient was, and the nature of the activity, though there would be a degree of degradation in the body irrespective of activity. The patient's weight and body shape would also have an impact on the forces being applied to the prosthesis.

193. One of DePuy's factual witnesses, Dr Brian Haas, was working as a specialist knee and hip orthopaedic surgeon at the Woodridge orthopaedic clinic in Colorado from 1990 until 2000, when he and a colleague founded the Colorado Clinic, which became a specialist referral centre. This clinic, rather like the tertiary centre in Oxford where Mr Whitwell works, treated difficult cases and cases that other orthopaedic surgeons did not wish to treat. Throughout the 1990s, Dr Haas was dealing on a weekly basis with cases of osteolysis requiring revision. I found him a frank and straightforward witness, and his evidence was largely unchallenged.
194. Dr Haas's experience was that osteolysis in MoP implants would not usually develop until the prosthesis had been implanted for 7 to 10 years, though he had seen cases in which there had been accelerated wear and osteolysis that occurred before 5 years. He said his experience of the incidence of failure due to osteolysis was shared by many other orthopaedic surgeons. It was a major focus at meetings of professional associations of surgeons at which the topic of polyethylene wear was discussed year after year. Patients with osteolysis were largely asymptomatic in the early years, and problems would only become apparent if they started to experience pain, or the prosthesis started to loosen. Mr Whitwell's evidence was consistent: osteolysis was something that developed over time, but if the hip was not subject to periodic X-ray follow-up, and it was well fixed, the consequences of the adverse reaction to the polyethylene debris would not become manifest until the hip began to loosen and the patient began to suffer symptoms. Professor Pandit and Mr Whitwell agreed that osteolysis may occur silently without any symptoms of pain or lack of function.
195. During the 1990s, once earlier problems experienced with the fixation of the prosthesis appeared to have been largely overcome, the search for alternative harder-wearing material combinations of the articulating surfaces of total hip implants became the subject of more intense scientific study in the orthopaedic community. Designers of implants began to focus on improvements to wear rates, so as to reduce the volume of wear debris. This led to a move away from conventional MoP towards alternative articulations. A third generation of CoC implants were developed and introduced during the 1990s, but concerns about potential fracture remained, and the size of the components meant that these prostheses were unsuitable for patients with smaller anatomies. At around the same time, experiments began with structural cross-linking of polyethylene through gamma or beta-ray irradiation, and the use of antioxidant agents such as Vitamin E to stabilise the material, producing early examples of cross-linked polyethylene.
196. It was also in the 1990s that MoM resurfacing prostheses were developed by Wagner and McMinn, using cemented and uncemented fixations. The advantage of the resurfacing technique was that it allowed for preservation of the femoral head and neck, with normal femoral loading, and reduced the risk of dislocation. That was one reason why hip resurfacing became a popular choice for younger and more active patients. Resurfacing arthroplasty started to be used in significant numbers in the mid-1990s. The "Birmingham" hip resurfacing was first clinically used in 1997.
197. In 1995, a conference took place in Santa Monica, California, which was attended by leading scientists, clinicians, engineers, academicians, industry-based scientists and engineers, and government regulators involved with orthopaedics, to discuss the future of hip prostheses and specifically the issue of alternative bearing surfaces to conventional polyethylene, focusing on MoM bearings. The conference brought



together people with clinical and scientific experience who had already undertaken research or had a specific interest in evaluating MoM bearing technology for total hip arthroplasty. The conference generated a “consensus statement” on the current state of knowledge in the industry, which was published in a leading orthopaedic journal, “Clinical Orthopaedics and Related Research”, in 1996.

198. The consensus statement referred to the fact that retrieval analyses had indicated that in the absence of specific design flaws, such as excessive clearance, the long-term wear rates of MoM hip prostheses were typically up to 20 times less than polyethylene wear rates. Despite noting that histological features seen with failed MoM components differed from those seen around MoP components, and express identification of the issues of metal ion and particle toxicity (and even potential carcinogenicity), the statement concluded that MoM technology was a worthwhile subject of further research and development as an alternative articulation, to address the problem of polyethylene wear particle-induced osteolysis.
199. The views of the orthopaedic community as to the need to improve on the longer-term survivorship of MoP prostheses because of the problems of particle-induced osteolysis were reinforced by the data from the National Swedish Hip Arthroplasty Registry (“SHAR”) published in a paper by Dr Malchau, one of the directors of the registry, in 2000. This report was usefully written in English. In this judgment, I shall adopt the same shorthand as was used throughout the trial and refer to Dr Malchau’s paper as “the SHAR 2000 report”; however, it is not to be confused with the annual report published by SHAR (in Swedish only), in that same year. SHAR, which was established in 1979, is the longest-established national joint registry. The data referred to the entire patient cohort (169,419 primary procedures and 13,561 revisions) and it was the only available data at the time on long-term survivorship rates. The orthopaedic community’s general view of survivorship of hip prostheses was largely informed by this registry. The SHAR data showed that younger and more active patients, particularly those aged under 55, were at greater risk of revision in all diagnostic groups.
200. Both Mr Whitwell and Professor Pandit confirmed that the received wisdom as to the likely survivorship of a hip prosthesis after a primary operation at the time when the Ultamet came onto the UK market was based on the track record of MoP prostheses recorded in the SHAR 2000 report. I consider the findings of SHAR in detail in Chapter 4 of this judgment (section 4.3) when evaluating the statistics. At this juncture, I note only that the Swedish data indicated that the survival rate of MoP prostheses declined rapidly after 10 years, especially in younger patients, a problem starkly illustrated by the graphs in the report. The rapid decline has been appropriately described as a “cliff edge” phenomenon. The SHAR 2000 report continued to drive the search for alternative, harder-wearing bearing surfaces.
201. Prior to the early years of the 21st Century, total hip arthroplasty was largely targeted at older patients (those aged 70 or over), with the purpose of restoring mobility around the home and the locality, and significantly reducing pain. The prevailing MoP articulation worked reasonably well in such patients by providing improved function and pain relief. Their activity levels tended to be low, and so the demands on the prostheses were relatively modest. The standard advice given to such patients was that such a prosthesis might last 10 to 15 years. In practice, as the evidence of Dr Hass illustrates, there would be those that failed earlier, including much earlier, and those

that would last much longer. The experience of clinicians was that the wear of the hip would increase with higher levels of activity, which reduced the length of time the prosthesis might be expected to survive in a more active patient.

202. There was a general reluctance in the orthopaedic community to proceed to replace the hip of a younger patient, because younger and more active patients appeared to be at greater risk of revision in all diagnostic groups. This was borne out by the data in the SHAR 2000 report. Mr Whitwell said that for younger patients who were sufficiently symptomatic, clinicians would still make the decision to replace; but they would generally try to keep the patient going with injections and/or physiotherapy until they fell within the elderly, less active group. He said that at that time, in such cases, orthopaedic surgeons lacked confidence that they had an implant which was going to outlive the patient.
203. Therefore, the major challenge for the orthopaedic community was to create an implant that was suitable for the younger, fitter patient who wanted to lead a more active life. They wished to develop prostheses to combat the problem of osteolysis and provide a greater range of motion whilst also reducing the incidence of dislocation. The larger sizes of femoral heads were developed to address the problem of dislocation and create greater stability.
204. One advantage of using metal as the liner material was that it enabled a larger sized femoral head to be used in the prosthesis. The maximum size of components in the bearing of the prosthesis is limited by the strength of the materials comprising the liner. Stronger materials can be thinner, and so accommodate a large femoral head within the same size acetabular cup. Metal liners allowed for the largest size heads with the thinnest liners. At all material times relevant to this litigation, a patient who received a 50mm acetabular cup could only have a 28mm or 32mm head size if they wished to have a polyethylene or ceramic liner, but they could have a 36mm head size with a MoM articulation. Over time, larger sizes have been developed for ceramic and HXLPE; this was an ongoing process throughout the 2000s.

### **3.2 THE DEVELOPMENT OF THE PINNACLE ULTAMET**

205. DePuy became part of the Johnson & Johnson group in August 1998. By then, Johnson & Johnson Orthopaedics had already embarked upon research into how to address the issues presented by using conventional polyethylene as a liner material. They were looking into ways of improving the durability of polyethylene, inter alia by the use of cross-linking techniques, but at the same time they were considering how ceramic materials might be developed as viable alternatives, and how to improve on the earlier designs of MoM articulations. Each of these projects progressed in parallel. Other manufacturers were carrying out similar research and development projects.
206. As a result, a new generation of alternative articulations to metal on conventional polyethylene were developed and launched at around the turn of the new century. These included the new generation of MoM designs, of which the Pinnacle Ultamet was one component, and the first versions of HXLPE which varied from producer to producer and developed over time. Subsequently these were joined by the current (fourth) generation of CoC articulations. The surgical, regulatory and manufacturing communities were all aware that these were all new products that had undergone *in vitro* testing. It was unknown how well they would perform in real life.

207. Some of the new generation of large head MoM articulations were initially developed to allow for the revision of failed resurfacing arthroplasties, though this was not the case with the Pinnacle Ultamet. That was specifically developed to give orthopaedic surgeons a complete range of choice of liner materials to use within the Pinnacle modular system. At least initially, it was the most popular choice within that system.
208. The Court had the benefit of evidence from some of the key individuals who were involved in the development of the new generation of MoM bearings at DePuy, including the Pinnacle Ultamet. DePuy's alternative bearings team was a multidisciplinary team comprising among others, surgeons, scientists, and engineers. Dr Haas, who was the surgical consultant on the MoM hip project, took the view that large diameter MoM components were a viable alternative to conventional MoP prostheses, which could address the problems of the dislocations and high revision rates for patients, especially those who were younger and more active. It was the evidence of Dr Haas and also that of Dr Richard Farrar, who was one of the senior project Bioengineers working within the research and development team on that project, that DePuy had much better manufacturing techniques than were available with the historic MoM prostheses. Therefore, they could manufacture the components with higher precision, better tolerances, and a better surface finish.
209. Dr Farrar's witness statement was not challenged. He was not directly involved in the design of the Ultamet, although he was involved in the development of an earlier product, a metal cup named Ultima. However, Dr Farrar's review of the publicly available literature and his identification of some of the major issues with the first generation of MoM products (forming part of his 2001 PhD thesis) had a significant influence on the steps taken by DePuy's research and development team to reduce the amount of wear released from the metal components in the prostheses to improve their performance.
210. Dr Frank Chan, who gave oral evidence, has a PhD in biomedical engineering from McGill University in Canada. His thesis was entitled "wear and lubrication of metal-metal bearings in total hip arthroplasty." He joined DePuy in 1999 as a senior scientist specialising in tribology (the study of friction, lubrication and wear). In that role, he was involved in the design and analysis of tribological systems for hip replacement applications. Until 2001, he was the lead tribologist on the Ultamet project and a member of the project core team. In 2001, he was appointed senior research engineer in biomechanical testing and analysis and moved away from tribology. In that new role, his responsibilities included evaluating the strength, mechanical properties and fatigue life of all types of orthopaedic medical devices. After further promotion within that department, Dr Chan left DePuy to work for another company in 2004.
211. I am satisfied on the evidence of Dr Chan and Dr Farrar that the Pinnacle modular system and the Ultamet liner were well-designed products, with low diametrical clearances between the metal femoral head and the bearing surface of the cup, aimed at producing low wear and maximum lubrication. It was obvious from the contemporaneous documents, and Dr Chan confirmed, that problems with increased friction in the 36mm heads observed during one specific test carried out during the design phase in November 2000, were due to a faulty component in the testing machine. Initially, DePuy had decided to use the same nominal clearances for the products using 28mm and 36mm heads, because their manufacturing processes were capable of producing implants with those low clearances. The results of that test led to

their decision to increase the clearance of the 36mm head product. Dr Chan explained that the difference between the intended specification and the new specification was marginal (four hundredths of a millimetre) and so, even after it was discovered that there was no problem after all, DePuy decided not to revert to the original specification.

212. The generic orthopaedic experts agreed that the rationale for the introduction of large head MoM implants was to reduce the incidence and extent of osteolysis and loosening and to improve survivorship in the young and/or active patient. There was the further potential benefit of improvement in range of movement of the hip joint, which potentially allowed an increased range of activity in the younger or more active patient. Additional reasons for the introduction of these implants included the improved stability of the hip joint (and therefore the reduction in the dislocation rate) which was potentially beneficial for the elderly patient as well. Surgeons saw the new generation of large head MoM articulations as attractive alternatives to the conventional MoP articulations, especially in younger and active patients, although there was no indication that their use should be confined to that cohort.

### **WARNINGS AND TECHNICAL INFORMATION**

213. DePuy produced technical documents to provide orthopaedic surgeons with information about their new products. Mr Whitwell confirmed that surgeons thinking of using a new product would obtain such material and read it before implanting the product for the first time; he explained that in practice, with the pressures of time in the operating theatre, a surgeon would be less likely to read the information and any warnings that came as part of the packaging of the products (the instructions for use or “IFU”s).
214. DePuy produced a Technical Monograph for the Ultamet in 2002 which expressly stated that the “*full biological response to metal particles or ions is currently unknown*”. It went on to raise the possibility of an immune response, mentioned the possibility of corrosion during phagocytosis and stated that “*in terms of toxicity, it was recognized that metal particles could cause a macrophage-mediated response with the production of various cytokines.*” It also noted that the response may be dependent on dose and cell type. It even went so far as to identify the theoretical possibility of cancerous tumours developing in response to metallic debris. It concluded “*overall, the biological effects of particulate or ionic metal... are somewhat uncertain.*”
215. Whichever way a surgeon went about informing him or herself of the risks before using a new product, it was incumbent upon the surgeon to have understood those risks before first implanting the device. Mr Whitwell confirmed that the orthopaedic community would not just read the technical literature – they would attend conferences and read general papers and publications and share information with their colleagues as well. He agreed that, irrespective of whether they read the IFUs or not, the information that was stated in the IFUs was something that an orthopaedic surgeon ought to know from the technical literature and other sources before ordering the product. He thought most of them would have known about the history of the MoM articulation, and the various problems with metal on conventional polyethylene in the past that had led to the view that there was a potential for improvement in performance and survivorship with the new generation of MoM prostheses.

216. Mr Whitwell agreed that an orthopaedic surgeon would understand that MoM articulations offered potentially, but not guaranteed, lower wear rates. As to whether they would be aware that it was possible that MoM bearings would release metal ions that could elicit an immune response, but that the full nature and extent of those reactions was unknown, Mr Whitwell said: “*you would hope the majority of surgeons would have picked that up in preparing for a new type of implant.*”

### **REGULATORY APPROVAL OF THE PINNACLE SYSTEM**

217. The Pinnacle system was approved for use in the USA and throughout the EU. In her witness statement, Ms Sally Hunter explained in some detail the process required to obtain a “CE Mark” in accordance with Council Directive 93/42/EEC (“the Medical Devices Directive”) which was implemented in the UK by the Medical Devices Regulations 2002, SI 2002/618. Medical devices could not be marketed in the EU without a CE Mark. Ms Hunter confirmed that the process for getting products certified in the EU involved a very rigorous exercise and required a high level of scrutiny both internally and by the Notified Body. In the UK that body was the British Standards Institution, BSI Assurance UK Ltd (“the BSI”) which is regulated by the Medicines and Healthcare products Regulatory Agency (“the MHRA”). The MHRA is an executive agency of the Department of Health and the UK’s regulator of, inter alia, medical devices. In France, where the Corail stem was manufactured, the relevant Notified Body was GMed.
218. When a device fell within Class IIb, as the Pinnacle acetabular system and its individual components initially did, the system involved a manufacturer demonstrating that the medical device and any associated products complied with the Essential Requirements set out in the Directive. It did so by compiling certain information and documents and having them available for review in a Technical File, which could be audited at any time by the Notified Body. In 2007, as a result of amendments to the Regulations, the Pinnacle system was reclassified as a Class III product and the process for compliance changed. DePuy had to submit a substantial body of documentation known as a Design Dossier to the BSI, and a CE mark would be awarded only if the BSI was satisfied (as it was) that the product met the Essential Requirements in the Directive.
219. The Design Dossier included, inter alia, a risk analysis, design verification (testing), and a review of technical and clinical information as well as information about its labelling and packaging. This included the IFUs which accompanied each component within the system. Mr Jonathan Cook identified in his witness statement the IFUs for the Pinnacle system and its components that were in use at various times. There were six different versions of the IFU relating to the Ultamet liner alone released between October 2002 and October 2010. The warnings in the IFUs included statements such as the following:

#### **For the Ultamet liner:**

*“Adverse Effects...*

*Histological reactions have been reported as an apparent response to exposure to a foreign material. The actual clinical significance of these reactions is unknown.*

*Implanted metal alloys release metallic ions into the body. In situations where bone cement is not used, higher ion release due to increased surface area of a porous coated prosthesis is possible...*

*Serious adverse effects may necessitate surgical intervention.”*

For the Corail stem:

*“Adverse events and complications...*

*The following are generally the most frequently encountered adverse events and complications in hip arthroplasty*

*Tissue reactions, osteolysis, and /or implant loosening caused by metallic corrosion, allergic reactions or the accumulation of polyethylene or metal wear debris or loose cement particles”*

A similar statement was made in the IFU for the Pinnacle cup.

220. The Essential Requirements set out detailed requirements in respect of, inter alia, product design (including chemical, physical and biological properties) and product labelling. The introductory paragraph notes that:

*“devices must be designed and manufactured in such a way that, when used under the condition and for the purposes intended, they will not compromise the clinical condition or the safety of patients, or the safety and health of users or, where applicable, other persons, provided that any risks which may be associated with their use constitute acceptable risks when weighed against the benefits to the patient and are compatible with a high level of protection of health and safety”.*

Paragraph 6 identified that: *“any undesirable side-effect must constitute an acceptable risk when weighed against the performances intended.”*

221. As required by the amended Regulations, the Pinnacle products were all re-certified under the Class III procedure by 1 September 2009. The Corail stem was a Class III product from its inception, and it obtained all relevant certifications in France.
222. DePuy’s quality system was accredited under ISO 9001/EH 46001 and the Directive. The ISO standard is an international standard that represents the requirements for a comprehensive quality management system for the design and manufacture of medical devices.

### **3.3 THE IMPACT OF THE INTRODUCTION OF THE NEW ARTICULATIONS**

223. Despite the initial rise in popularity of the new generation of MoM articulations, Table 3.4 of the 2016 Annual Report of the National Joint Registry for England, Wales, Northern Ireland and the Isle of Man (“NJR”) referred to in paragraph 53 of Mr Whitwell’s Expert Report, indicates that for most of the period from 1 April 2003 to 31 December 2015, MoP was still the most popular articulation for total hip arthroplasties, irrespective of the means of fixation. The vast majority of cemented implants have MoP articulations. For uncemented implants, MoP was overtaken in popularity by the latest generation of CoC articulations in 2010, though only for a

short time. After reaching a peak in 2011 the use of CoC started to decrease, whilst MoP articulations again became predominant. However, that table does not draw a distinction between conventional polyethylene and the various forms of HXLPE that were gradually being introduced onto the market throughout the period in question.

224. The use of uncemented large head MoM prostheses such as the Pinnacle Ultamet increased rapidly in the early part of that decade and peaked in 2007, before starting to decline in 2009, sharply dropping off in 2010 and thereafter reducing to an almost negligible number, for reasons which I shall go on to explain. Very few Pinnacle MoM articulations were implanted in 2011-2012.
225. The Claimants' engineering expert Professor Gill described HXLPE as a "*game changer*". Mr Whitwell preferred to use the phrase "*significant improvement*" when comparing it with conventional polyethylene. He agreed that studies indicated the initial results for HXLPE were largely encouraging in terms of survivorship, with a reduction in wear rates. He referred to a study carried out by a colleague of his in Oxford, Dr Glyn Jones, on one of the specific HXLPE products, which gave very encouraging ten-year data in terms of very low volumetric and linear wear rates measured in vivo. The study stated that there was no evidence as yet that these findings will result in lower revision rates or improved functional outcomes (in the longer term), but, whilst stressing the need for caution, Mr Whitwell said that as an orthopaedic community "*we are hopeful*." He accepted that there may be different long-term performances for the different types of HXLPE that had been introduced on the market over the past 17 years. The orthopaedic community did not yet know the long-term benefits of the stabilisation techniques used on the material, and he was aware that some studies suggested that HXLPE has a higher inflammatory potential than conventional polyethylene.

### **3.4 EVENTS LEADING TO THE WITHDRAWAL OF THE PINNACLE ULTAMET FROM THE MARKET**

226. The awareness in the orthopaedic community that there were potential problems with large head MoM hip implants was gradual. The first problems that were identified were with hip resurfacing components, rather than with total hip arthroplasties. Mr Whitwell's own experience since becoming a consultant in 2005 largely related to resurfacing arthroplasties. He was involved in the Oxford Hip Research Group that published widely on the field of MoM arthroplasty, after operating on one of the first patients recognised to be suffering from ARMD in 2005 – that patient had a resurfacing arthroplasty. It was in the resurfacing context that ARMD was first reported in a national orthopaedic conference in 2008. Over the next few years, scientific publications in peer-reviewed journals helped to increase awareness of it in the orthopaedic community.
227. During 2008, articles appeared in newspapers that highlighted failure rates, and also drew attention to possible side-effects from the release of metal ions. From around then onwards, media reporting started to describe concerns about MoM hip resurfacings and total hip arthroplasties in an increasingly sensationalist manner that would increase the likelihood of a reader becoming concerned or anxious if he or she had had such an implant. The fact that the overwhelming majority of patients with MoM total hip arthroplasties have well-functioning hips (as the MHRA acknowledges on its website) is something that the media generally failed to mention.

228. On 8 March 2010 DePuy issued an urgent field safety notice in respect of its ASR resurfacing implants, prompted by a higher than expected revision rate in patients who had such implants with femoral head sizes less than 50mm. A few weeks after this, on 22 April 2010, the MHRA issued specific guidance relating to all MoM hip replacements following reports of revisions of such hip replacements involving soft tissue reactions associated with hip pain. The guidance made no distinction between hip resurfacing arthroplasties and total hip arthroplasties. It advised that patients be followed up at five years post-operatively and more frequently in the presence of symptoms. Patients with painful MoM hip replacements should be investigated, specifically tested for Co and Cr ion levels in the blood and given an MRI or ultrasound scan. If either Co or Cr ions were elevated above 7 parts per billion (ppb), then a second test should be carried out in three months. If the imaging revealed soft tissue reactions, fluid collections or tissue masses, revision surgery should be considered. The purpose of this guidance was to enable those patients who had ARMD to be identified sufficiently early to enable them to undergo revision surgery at a time when there was the best opportunity to produce a good result.
229. In August 2010, DePuy issued a further Urgent Field Safety Notice indicating that it had voluntarily decided to recall all ASR products because of the reported higher than expected short-term revision rates. This led to the MHRA issuing a specific notice about the recall which warned that surgeons should not implant DePuy ASR products.
230. Media reporting became even more sensationalist in the wake of the withdrawal of the ASR resurfacing device. Again, no attempt was made to distinguish between resurfacing implants and those used for total hip arthroplasties. Phrases such as “*toxic hip implants*” were used in the tabloid press, and it was suggested in some articles that there were systemic problems, and that metal was poisoning the bodies of patients. Typically, the absence of any scientific research to support the alarmist contentions in the articles or reports would be mentioned, if at all, in the last paragraph. Television news and websites also carried stories about the possible dangers of metal on metal hip replacements, including a joint investigation by Newsnight and the British Medical Journal which referred to “*poorly regulated and potentially dangerous*” hip devices, and suggested that wear debris might be carcinogenic.
231. The joint report of the expert behavioural psychologists, Dr Rubin and Dr Horne, whose evidence I was invited to read, indicated that media reports of possible adverse consequences from MoM prostheses had the potential to cause a so-called “nocebo” effect in some patients, whereby the expectation of symptoms becomes self-fulfilling. This could manifest itself in heightened anxiety, causing patients to monitor themselves more closely, and in patients noticing and remembering symptoms that they might otherwise have missed or disregarded. It could also lead to a subjective perception of increased pain and/or a loss of function; a misattribution of pain; and/or an increased willingness on the part of some patients to seek help, or to consent to revision surgery.
232. These matters also had some impact on the willingness of surgeons to recommend revision surgery. This conclusion is not just supported by the evidence of the behavioural psychologists. It was the evidence of the generic orthopaedic experts, borne out by the evidence of their colleagues who gave factual and expert evidence in the test cases, that symptoms and reports of pain were of considerable importance to the decision whether to monitor a patient in accordance with the MHRA guidance and



to the decision whether to revise. Mr Whitwell, for example, made it clear that pain was a major factor in his decision-making. If a patient with a MoM hip had symptoms, he would revise.

233. Accordingly, if a patient developed somatic pain as a result of reading the adverse media coverage, and reported it, it might well be enough to convince an orthopaedic surgeon, particularly an ultra-cautious one, that there was sufficient cause for concern to revise their MoM prosthesis. As Mr Whitwell pointed out, extensive soft tissue necrosis presents specific challenges for the surgeon which can be much more difficult to overcome than damage to the bone, particularly if muscle damage occurs. Knowledge of this is bound to have affected surgical attitudes and made some consultants more willing to revise than they might have been if, for example, the problem was potential osteolysis.
234. The expert psychologists agreed that the sensationalist media reports probably did have some effect both in increasing the rate of, and accelerating the timing of MoM hip implant revisions, though they were unable to express a view on the degree to which this impacted on the reported revision rates. They very fairly pointed out that any tendency towards increased rates of revision may have been opposed by the availability of objective measures and guidance to apply these before deciding to revise. However, the problem was that there were no clear diagnostic criteria for ARMD. The MHRA guidance suggested that the results of cross-sectional imaging would influence the decision to consider revision, but it shed no light on how the consultant should approach a case in which that imaging revealed nothing of concern, other than by continuing to monitor the patient annually. Much would depend on the attitude of the treating clinician, and the extent to which he or she placed emphasis on the subjective experience of the patient.
235. Mr Whitwell agreed that there would have been some revisions undertaken, or undertaken earlier, as a result of concerns engendered by the media, and that the number could not be quantified. At least one of the lead claims falls into this category. Professor Pandit's evidence of the impact on the surgical community of reports of systemic problems with MoM prostheses was to similar effect. He referred to the self-fulfilling prophecy whereby panic engendered by the sensationalist media reporting increased the revision rates, and higher revision rates, once known, cast further doubt on the prosthesis, and thus prompted surgeons to offer revision more readily and patients to accept revisions more readily.
236. As the evidence in the lead claims demonstrated, the 2010 MHRA guidance was not consistently followed. In practice, some surgeons (or centres) adopted a more conservative approach and used lower thresholds for surveillance and surgical intervention. There were some centres around the country in which all patients with MoM prostheses were tested for blood metal ion levels, irrespective of whether they were symptomatic. There were patients who were subjected to more extensive investigations even if their Co or Cr levels were less than the MHRA surveillance threshold of 7ppb. Local guidelines in certain areas adopted a lower threshold of 4ppb, and some orthopaedic consultants took the view that even that threshold was too high. Four of the surgeons who performed revisions for the six lead claimants used a lower threshold for investigation and revision than the MHRA guidance.

237. Even Mr Whitwell adopted a more cautious approach than that recommended by the MHRA. His view could be summed up as “if in doubt, revise”, and that included the situation in which the imaging showed a collection round the trochanter. He was only prepared to leave and monitor the smallest of collections of around 5cm<sup>3</sup> and below, and then only if the wall of the collection was thin. Given that Mr Whitwell worked in a tertiary centre and therefore saw the worst types of case, his unwillingness to wait and see how things developed is understandable. His attitude was also based on his personal experience that when imaging revealed a collection, the patient was not truly asymptomatic because even if they initially denied symptoms, close questioning would normally reveal some evidence of pain or discomfort which they had grown to live with. Mr Whitwell was far from alone in adopting this ultra-cautious approach. Indeed, it was an approach that Mr Whitwell shared with colleagues to whom he gave lectures, and is likely to have influenced their thinking.
238. The strong impression that I received from the evidence, particularly in relation to the lead claims, was that once awareness of the phenomenon of ARMD spread, the orthopaedic community tended to act out of an abundance of caution. Ironically, the design of the Pinnacle modular system which made it relatively straightforward to swap the head and/or liner had a part to play in this. There were undoubtedly some surgeons who would investigate and/or revise a patient merely because he or she had a MoM hip and some unexplained pain: Mr Herlekar, who revised the hip of the lead claimant Mr Woods, was one example. Others, faced with the possibility that ARMD might develop, decided to swap the head and/or liner on the basis that it was better to be safe than sorry. Yet in both types of scenario, the patient’s hip may well have lasted for more than 10 years, without the patient experiencing any adverse soft tissue damage. Although this ultra-conservative approach will have resulted in some patients who had ARMD being diagnosed earlier than they otherwise would have been, it also resulted in revisions being carried out on some patients who were not suffering from ARMD at all.
239. On 2 March 2011 the British Orthopaedic Association (“BOA”) made an announcement concerning large bearing MoM total hip replacements. It reported that at the 2011 British Hip Society Annual Conference, several units reported survivorship data for these devices which demonstrated poor survivorship in the short to medium term. It stated that there was a predominance of the ASR XL device, which by then had been withdrawn, but that large diameter MoM devices from other manufacturers *may be* showing similar results. In the light of this data, the BOA advised that the existing MHRA guidance should continue to be followed and the use of such bearings in primary total hip replacement should be carefully considered and possibly avoided.
240. On 28 February 2012, the MHRA issued revised guidance on management and monitoring of patients with existing metal on metal hip replacements. It identified four groups of patients who should be followed up as per the previous 2010 advice, whether symptomatic or asymptomatic. These included patients with MoM total hip replacements with a head diameter of 36mm or above.
241. There is no current indication for the use of large head MoM prostheses in total hip arthroplasties with any patient group. That has been the case since March 2012, when the British Hip Society issued a statement advising that large diameter MoM primary total hip arthroplasties using bearings of 36mm or above should no longer be

performed until more evidence was available, except in properly conducted and ethically approved research studies. In May 2013 DePuy announced that with effect from August 2013 it would be discontinuing the Pinnacle Ultamet MoM articulation worldwide.

242. In their joint statement, the generic orthopaedic experts stated that it remains an accepted consensus among orthopaedic surgeons that all large head MoM implants are associated with an increased risk of early failure, but most currently remain satisfactory under long term surveillance. Mr Whitwell said that when the MHRA guidance first came out, there were specific clinics run in Oxford to monitor the patients, but as the understanding of the ARMD phenomenon in the orthopaedic community improved, patients with MoM prostheses were brought back into routine clinics for follow-up. He and his colleagues would now only intensively monitor high risk patients, (which definition included any person with a large head MoM prosthesis), and that monitoring would be annual.

## **CHAPTER 4: THE CLAIMANTS' CASE**

243. Surgeons and patients alike at the material time believed that the new generation of MoM hip prostheses would survive for longer than the existing, predominantly MoP prostheses. That was also the aspiration of DePuy. They were designed with that objective in mind. Many, if not most of the Pinnacle Ultamet prostheses achieved it. Whilst it was hoped that the new generation of hip implants would be an improvement on the existing prostheses in this and various other respects, when they were launched on the market it was not known how well they would perform in real life. A failure to meet aspirations does not equate to a lack of safety.
244. The possibility that some patients might suffer an adverse histological reaction which could lead to clinical consequences and revision surgery was flagged up in the IFUs and technical brochures, together with the fact that the extent that the risk might materialise was unknown. Orthopaedic surgeons should have been aware of that inherent risk and would have been if they read the literature that was disseminated by DePuy. The regulators were also aware of it. The key issue is whether the manifestation of the known risk, in terms of early failure of the prosthesis, was sufficient to bring the Pinnacle Ultamet product below the objective safety threshold laid down by the Act.

### **4.1 REVISION RATES AS AN OUTCOME MEASURE OF SAFETY**

245. The claimants submitted that revision rates are an appropriate outcome measure by reference to which the Court can assess the safety of a prosthesis for the purposes of determining if it met the entitled expectation of safety required by s.3 of the Act. Revision, in this context, means a hip replacement procedure that involves the removal or replacement of at least one component of the implant, and thus includes the exchange of a head and/or liner in a modular system such as Pinnacle. They contended that the need for revision is a reliable and well-defined outcome measure that (subject to certain caveats) is a direct reflection of implant failure which is capable of being compared between different implant types.
246. The claimants accepted that revision as an outcome measure will not capture all aspects of an implant's performance that impact on safety. Some patients with failed

implants may refuse revision or be too unwell for surgery; some revisions may take place for reasons other than implant failure, such as infection, mal-alignment, dislocation or fracture. So not all revisions are failures, and not all failures are revised. However, they contended that it is appropriate to use revision as a proxy for implant failure, because (a) the incidence of revision for other causes is very low and (b) it is safe to assume that the incidence of revision for other causes and failure to revise because of health factors or refusal by the patient is broadly the same across all categories of prosthesis.

247. The claimants realistically did not put their case on the basis that the public was entitled to expect a MoM hip replacement to have the same survival rates as, or better survival rates than, alternative articulating surfaces that were then available. Their case is that, at the time when it was launched on the market, the Pinnacle Ultamet 36mm head prosthesis carried with it a materially increased risk of failure/early revision than relevant comparator implants (because of the propensity to cause ARMD) and that is why it fell below the standard of safety that the public generally were entitled to expect.
248. Mr Oppenheim described the entitled expectation as being one of “comparable safety” with all other products that were then available on the market. That expression has the potential to confuse; “comparable safety” would normally be understood to mean the same or better. That is very different from an expectation that a new prosthesis will not carry with it a *materially increased risk* of early failure. It is the latter formulation that reflects the pleaded case of an abnormal potential for damage. On the claimants’ case it is not enough to demonstrate that the rates of revision for an Ultamet prosthesis were higher than for the comparator, they must establish that they were very much higher, and that the higher rates of revision are a reliable measure of a failure to meet the safety standard that the public was entitled to expect of a prosthesis in 2002.
249. DePuy submitted that survivorship is only one of the features against which the performance of a prosthesis can be judged. The functional performance of a prosthesis and the ability of the patient to return to normal activity is important, and the ease of revision in a system such as the Pinnacle is a beneficial feature, especially for a younger patient. So too is the improved stability that a large MoM head provided. They pointed to the agreement by the generic orthopaedic experts in their joint statement that when the new generation of MoM products was introduced there was an expectation of potential benefits or improvement in various respects, in comparison with the products that were already in use. Mr Whitwell acknowledged that the Pinnacle Ultamet achieved a range of such benefits, apart from the concern over its short-term survivorship. Some of the claimants in the lead claims gave evidence about how well their new Pinnacle hips performed and how happy with them they were, at least for the first few years.
250. I accept that if a new prosthesis afforded greater stability and range of movement and could be revised more easily than existing models, as was true of the 36mm head Pinnacle Ultamet prosthesis, it would have a range of benefits that would make it attractive, especially to the younger and more active patient who wished to be able to engage in sporting and other activities that might be expected to cause greater wear to the implant. However, whilst the public generally might be expected to accept *some* increased risk of early failure of the implant in those circumstances, bearing in mind those benefits, it cannot seriously be suggested that they should be expected to accept

a much greater risk of early failure, if that is indeed what the statistics reliably demonstrate. That was Mr Whitwell's view, and I agree with him.

**THE USE OF A 10 YEAR PERIOD TO MEASURE THE RATES OF REVISION**

251. The claimants relied on comparative cumulative risks of revision ("CRR") to demonstrate that a materially increased risk of early failure existed at the time when the product was introduced to the UK market in 2002. By early failure, the claimants meant revision within 10 years of primary surgery, a definition adopted by Mr Whitwell, though that was not a generally adopted measure in the orthopaedic community. Professor Pandit, for example, defined early failure as failure within the first 2 years of implant.
252. DePuy criticised the selection of a 10-year period as a measure of safety on the basis that it was artificial and dictated by the limitations of the available evidence at the time of trial. Since the first Pinnacle MoM hips were implanted in 2002 before the NJR existed and the data on which the claimants rely began to be collected, there would be hardly any data indicating their actual survivorship levels beyond 10 years which, DePuy said, would have some bearing on whether they out-performed the existing prostheses that they were designed to improve upon. The claimants' case is based on statistical probabilities. On the face of it, there is some force in that objection: if, for example, 20% of metal on conventional polyethylene prostheses failed within 15 years of implantation, but only 10% of MoM prostheses failed within the same period, the latter would appear to be a considerable improvement on the former, and the comparative performances at 10 years would only reveal part of the picture.
253. However, I must remind myself that the focus is on the entitled expectation of *safety* of a prosthesis in 2002. At the time when the Ultamet was introduced, it was generally expected that a hip prosthesis would last for about 10 years, at least, before requiring revision, though it was also expected that there would be some incidence of earlier failure, and that the incidence would be greater in younger and more active patients. The expectation of 10-year survivorship was reflected in the advice given to patients, and it was based on clinical experience and on data from SHAR relating to the performance of MoP prostheses. The 10-year period was the period that the orthopaedic community were trying to improve upon by developing alternative, harder wearing materials than conventional polythene for use in the bearings. It was also the period of survivorship used by the National Institute of Clinical Excellence ("NICE") over which the performance of a hip prosthesis was measured, in order to set a benchmark for the best performing products that NHS consultants used as a reference for the selection of implants.
254. It is also an important part of the background that surgeons were very reluctant to implant younger and more active patients with prostheses that they believed would last for at least 10 years in an elderly inactive patient, because they wanted to avoid revision or re-revision during the patient's remaining lifetime. If they would not use the existing prostheses on such patients unless they felt they had no choice but to carry out an arthroplasty, they would not have used an alternative implant in any patient if the alternative was far more likely than the existing ones to fail within 10 years. That would defeat the very objective they were trying to achieve, even if in the

long term, those which did not fail early would last for much longer than the comparators.

255. Mr Whitwell said that he did not regard the prospect of longer term survivorship as a trade worth making even for the younger patient. The soundness of that assessment may be a matter of fact and degree, because a slightly higher risk of failure in the short term might well be objectively regarded as a trade worth making if the alternative were enduring many years of pain whilst waiting to reach the age where a consultant was willing to perform primary surgery. However, the claimants' case is postulated on there being a materially higher risk of early revision. The prospect that the CRR for the different prostheses might converge at some later period, for example, after 15 or 20 years, which was the subject of some debate at trial, has no bearing on the issue of defect when the defect is legitimately defined as an abnormal propensity for early failure. Although I accept that there is a degree of artificiality about using a 10-year period, I consider there is a sufficient justification for it.
256. One advantage of using revision as an outcome measure is that it is not susceptible to the subjectivity of patient-reported measures. However, the decision to revise still involves a degree of subjectivity. The data in the NJR on which the claimants relied are informed by the views of the operating surgeon as to the need and reason for the revision. If and insofar as the decision was taken to revise for ARMD, there is and was no accepted consensus about the factors that would lead to that diagnosis. Moreover, when faced with particular symptoms or test results, one clinician might decide to wait and monitor the patient, whereas another might take a more precautionary approach and decide to operate straight away, so as to avoid the prospect of worse damage arising later.
257. Mr Oppenheim sought to make great play of the fact that Professor Pandit accepted in cross-examination that if all that was known about the Pinnacle Ultamet by 2012 had been known in 2002, it would have had no indication for use in 2002. However, the Court is not concerned with the performance of the Pinnacle prosthesis, but with its *safety*. These are not synonymous. It is quite clear from the evidence that I have seen and heard in this trial that the incidence of early revision of MoM prostheses recorded or predicted in the NJR was something that nobody anticipated – indeed, given the state of scientific knowledge then (and now) it was unpredictable. If that information had been known in 2002 I agree that it is likely that surgeons would not have chosen the Pinnacle Ultamet in preference to one of the other types of articulation, as it would not have been regarded as sufficiently clinically effective; but it does not necessarily follow from this that the Ultamet was defective within the meaning of the Act.
258. The fact that in the short-term a prosthesis did not perform as well as the competitors it was designed to improve upon, will not necessarily demonstrate that it failed to meet the entitled expectation of safety. The forensic point that Mr Oppenheim was making only has force if the reported incidence of revision was a reliable indication that the propensity of the Pinnacle MoM prosthesis to fail within 10 years for ARMD was materially greater than that of an appropriate comparator. That depends on the reliability of the statistical evidence, to which I now turn.

**THE EXPERT EVIDENCE ON STATISTICS AND EPIDEMIOLOGY**

259. In support of their case, the claimants called expert evidence in the fields of biostatistics (Ms Alison Smith) and epidemiology (Professor MacGregor). Ms Smith was the lead statistician for the NJR from 2009 to 2012. She undertook the survivorship analysis for Part 3 of the NJR's Annual Reports for 2011 and 2012 which were the first to specifically compare MoM bearing surfaces with other types of articulation. Both were well suited to give evidence as experts on these topics and were generally fair and balanced, though there were some aspects of their evidence that I found less satisfactory than others.
260. Shortly after Ms Smith and Professor MacGregor had finished giving their evidence, DePuy indicated that they no longer intended to call their experts in those fields, Professor Scharfstein and Professor Hutton. On the day when Professor Scharfstein had been due to give his evidence, I convened a case management hearing and invited submissions on the effect of this unexpected development, and how the Court should proceed. It was common ground that no reliance could be placed on the reports of the experts who were not called, or on any contributions they had made to the joint expert report, save where there was agreement between them and the claimants' experts. The Court must also discount any evidence of experts in other fields that was premised on the Court's acceptance of any contentious evidence of those experts. That is the approach I have adopted.
261. In their final submissions, the claimants submitted that the Court should approach with caution any evidence given by Ms Smith or Professor MacGregor in cross-examination in which they agreed with points put to them that had been made by Professor Scharfstein or Professor Hutton, which might have been tested with those witnesses in cross-examination had they been called. I see no good reason to do that; both experts had no problem in qualifying their answers (and did so, if they felt it necessary). They made it clear if they disagreed with the premise upon which the questions were put, or with the methodology that had been adopted by the defence experts. If they agreed with a proposition that was put to them in cross-examination, or accepted the methodology or data used or compiled by the opposing experts that was put to them, that is their evidence, and I will evaluate it in the same way as any other evidence they gave.
262. The claimants also complained that re-examination of both witnesses was attenuated given the time pressure to complete their expert evidence in accordance with the agreed trial timetable, and that their examination in chief and re-examination would have proceeded on a different basis had it been known that the defence experts were not going to be called. Ms Smith and Professor MacGregor would have been re-examined on points that Mr Oppenheim had decided to make instead by way of cross-examination of the defence experts. The claimants therefore sought to make submissions on matters that they said would have been put to DePuy's experts had they been called, where it was relevant and appropriate to do so.
263. The Court must ensure that all parties to litigation are treated equally fairly. However, the trial timetable was set and agreed well in advance, and an equal time was set aside for the expert evidence on these topics to be given by each party's respective experts. DePuy were under no obligation to call their own experts. The claimants' legal team should have been prepared to meet that eventuality, however unlikely they thought it

was to happen. I would not have allowed extensive examination in chief in the light of the substantial materials that the experts had already placed before the Court. In any event, if this development had truly caused a serious disadvantage in terms of a topic being insufficiently addressed, for example, I would have expected Mr Oppenheim to have taken the opportunity to ask the Court for permission to recall Ms Smith or Professor MacGregor. Since they would have been present during any cross-examination of Professor Scharfstein and Professor Hutton, they would have been available for recall on the days set aside for DePuy's expert evidence. No such application was made.

264. This trial proceeded on the pragmatic basis that neither party was expected to put every aspect of their case to every relevant witness. Bearing in mind how matters developed, I am prepared to allow the claimants a substantial degree of latitude. However, I am not prepared to allow them to introduce statistical material that was neither included in their expert reports nor addressed orally by their expert witnesses, on the basis that they had intended to put it to a defence expert in cross-examination, as that would be unfair to DePuy.
265. This is of some importance, because in the course of closing submissions, Mr Oppenheim sought to rely upon a calculation of how the magnitude of the difference between the CRR for the Pinnacle Ultamet prosthesis and comparator prostheses recorded by the NJR would be affected if the Court were to conclude that all the cumulative confounding factors inflated by 50% (or some lower percentage) the figures for revisions of the Pinnacle Ultamet prosthesis that were recorded during the data period. This calculation was not part of any expert report or the subject of any expert evidence. Mr Oppenheim said the methodology used was the same as the methodology used in a calculation put in cross-examination to Professor Pandit as a criticism of the analysis of the statistical evidence in a paper by a team headed by Professor Matharu published in December 2016, to which Professor Pandit was a contributing author. He submitted that no issue had been taken with Ms Smith in cross-examination as to that methodology, and that the exercise was an appropriate one. The point was being made on an illustrative basis.
266. However, Ms Smith never gave any evidence of relevance. She did give some evidence about that Matharu paper. She confirmed that the only criticism of it articulated in her supplementary expert report was that the authors had interpreted ARPD in the NJR data as meaning ARMD (which was not, in fact, a valid criticism, as that is how the surgeons filling in the forms interpreted ARPD). However, she said she also had some concerns about the way in which the authors of the paper had carried out some of the analysis of the data, chiefly because the number of revisions should have been divided by the total length of time observed. She drew those concerns to the attention of the claimants' legal team, which resulted in them being put to Professor Pandit in cross-examination. Professor Pandit, who is not a statistician, and was not responsible for the statistical analysis in the Matharu paper, was in no position to quarrel with the points that were being put to him. He largely accepted the criticism, though he quite fairly observed that the paper had been approved by the NJR.
267. As to the extent to which confounding factors would have to impact on the underlying data to explain the difference in the CRR relied on by the claimants, Ms Smith simply said (in the context of considering the impact of the 2010 MHRA guidance and the



lowering of the threshold for intervention) that it would have to have increased the reported revision rates on a “*very large scale*” to explain that difference. She did not attempt to quantify the impact on the data of the confounding factors she identified or accepted, save in re-examination when she said that it seemed implausible to her as a statistician that 42% or more of revisions fell into the category of having no proper clinical justification. She got that figure of 42% from a comparison of five-year data for patients whose primary operations occurred prior to April 2010 and five-year data for patients whose primary operations occurred after that date, even though she herself sounded a cautionary note about the reliability of the pre-2010 data.

268. Even taking that 42% figure at face value, that is the nearest that the claimants ever came to adducing any statistical evidence about the magnitude of the impact that any one or more confounding factors would have to have on the data in order to significantly reduce the differences between the comparative CRR on which they relied. In those circumstances, I am not prepared to place any reliance upon a document that was handed up during final submissions, which has not been prepared or adopted by an expert statistician and tested in cross-examination. The point that the claimants wished to make is one that could and should have been made through Ms Smith.

**USE OF HINDSIGHT TO DENOTE THE ENTITLED EXPECTATION OF SAFETY**

269. There was some debate as to the propriety of using data that came to light *after* the product came onto the market to inform the entitled expectation, as opposed to the question whether that expectation was met. The issue arose because the claimants’ case necessarily involves reasoning backwards from what they contend are reliable statistics on the 10-year CRR for the Pinnacle Ultamet and a comparator prosthesis or prostheses, submitting that there was a material difference in the risk of early revision for the Ultamet and concluding that therefore it did not meet the entitled expectation of safety.
270. An important factor in the evaluation of the entitled expectation would be the level of risk of early failure that was regarded at the relevant time as being acceptable, or tolerated for a comparator prosthesis, even if attempts were being made to improve upon it. This was the degree of risk of early failure that any patient offered a total hip arthroplasty in 2002 could reasonably be expected to accept and would in fact accept if they went ahead with the operation. If the Pinnacle Ultamet had a much greater risk of early failure than that, then it might easily be described as falling below the entitled expectation of safety. The claimants would not need to look at subsequent information or to use hindsight to establish the proposition that a person would be entitled to expect that a new prosthesis coming on to the market in 2002 would not carry with it a risk of early failure (necessitating revision surgery) that was far in excess of what was believed at the time to be the risk of early failure for an established prosthesis that he might otherwise be offered, if he was offered one at all.
271. I have far greater difficulty in accepting the proposition that if, unbeknown to anyone at the time, the risk of early failure of an alternative prosthesis was much lower than everyone believed, that lower risk should condition the entitled expectation of safety, but that is the claimants’ case. They contend that a prosthesis would not meet the entitled expectation of safety if its risk of early revision was materially greater than the CRR of a comparator prosthesis, even if, on the basis of all the information that

was then available, patients and surgeons expected a much higher incidence of early failure with the alternative prosthesis, and their advice to patients would have been premised on or, at the very least, conditioned by that expectation.

272. Mr Oppenheim submitted that a person would be entitled to expect that a new product would not have a materially increased CRR by comparison with the entire class of available alternative products, irrespective of whether the CRR for the latter were known or unknown, and irrespective of what the actual revision rates were. That proposition, he submitted, would not involve any illegitimate use of hindsight or retrospection. What conditions the entitled expectation of safety is the state of the art in terms of safety for the relevant range of alternative implants for use at the time the product was on the market, whether known or unknown. That would include any new products that were available at the same time as well as established products (a proposition that, save for the intra-Pinnacle comparison, went beyond the claimants' pleaded case).
273. Whilst I agree that there must be a minimum safety threshold that applies across the board to all products, new and old, in my judgment Mr Oppenheim's formulation sets the bar too high. In principle, what the public generally was entitled to expect in terms of safety at the time the product went onto the market cannot be measured by reference to information about the performance of a comparator product that only came to light subsequently. The Court is evaluating an acceptable level of safety at the relevant time, and that evaluation must be conditioned by the circumstances that were then known (including any published information about the product, and any relevant safety standards or regulations). Whilst the entitled expectation of safety is not the same as the actual expectation, that does not mean that it can be equated with what *might* (or might not) have been regarded as an appropriate safety standard, if the public had known more than it did.
274. Depending on the nature of the product and the material circumstances, the entitled expectation of safety may be higher or lower than the actual safety record of a comparator product. A comparator may turn out to have a particularly poor or a particularly good safety record, and what the public generally would be entitled to expect in terms of safety could be somewhere in between. Although in theory actual expectations may differ from the entitled expectation, that will depend on the relevant circumstances. In my judgment, the public would not be entitled to expect a new prosthesis to have a much lower failure rate in the first ten years after primary surgery than anyone would have regarded as acceptable at the time when it came onto the market, just because in the light of hindsight it could be demonstrated that an existing prosthesis performed better than anticipated.

### **MATERIALITY**

275. It was impossible to glean from the claimants (or their experts) what parameters they said made a material difference statistically, so as to enable an abnormal propensity for early failure to be inferred. Ms Smith agreed that where variations in statistics (between comparator prostheses) are relied on, "*strength in the sense of magnitude*" was important. In cross-examination she agreed that statistically, the greater the magnitude of difference in revision rates between bearing surfaces drawn from reliable data, the more likely it is that there is a genuine difference, and that a large relative risk would be stronger than one that is just slightly bigger than the

comparison. She added, very fairly, that absolute effect size is probably even more important than relative risk because with relative risk “*you could be twice as likely, for example, but the overall risk is small, so twice as likely for an overall risk isn’t necessarily a big effect size, so the absolute effect size is important too.*”

276. Mr Whitwell said in his expert report that the surgeon carrying out a total hip arthroplasty in 2002 “... *was entitled to expect that the risk of early revision for a MoM articulation would not be materially higher than the risk of early revision for relevant alternative bearing surfaces available within the same modular system, such as MoP CoP and CoC.*” He was asked in cross-examination to explain what he meant by “materially higher”, but he appeared to have considerable difficulty in answering the question. Ultimately it appeared that the point he was trying to make was that surgeons would not have expected to see the actual revision rates that were recorded in the NJR data. That does not really assist the Court in ascertaining at what point the claimants contend the variance in the risk of revision becomes material for the purposes of assessing safety. That question was never given an answer, let alone a satisfactory one.
277. In their final written submissions, the claimants’ counsel relegated their contentions on this issue to a footnote, in which they said that “*the increased risk of early revision clearly falls within any common sense judicial understanding of what would constitute a materially increased risk. Materiality is a matter for judicial assessment in this domain, as it would be in the field of informed consent.*” That is just an elegant way of avoiding giving an answer.
278. As Mr Antelme pointed out, materiality must be evaluated in the context of the particular circumstances with which the Court is concerned, and so it is of little help to consider what might be regarded as material in a wholly different context, such as informed consent. Depending on the circumstances, a new pharmaceutical product which carries with it a 5% greater chance of killing or causing serious injury to a patient than an existing product prescribed to treat the same disease *might* be regarded as falling below the entitled expectation of safety. By contrast, a hip prosthesis that has a 5% CRR within 10 years might not fall below that standard even if another hip prosthesis has only a 3% or a 2.5% CRR within the same period. That example illustrates Ms Smith’s point about the absolute effect being even more important than the size of the relative risk. Nor is it an answer to say, as Mr Oppenheim did, that if there is a 3 to 5 times difference in the revision risk, that is plainly material. Even if it did not matter to the outcome of this litigation where the line is to be drawn, it may matter a great deal in other cases.
279. Mr Antelme submitted that it would be inappropriate to find that a risk was materially increased if it was less than double the risk inherent in the comparator product, though a finding that it is double or more would not be the end of the analysis. He relied upon the approach in Australia as described in the 2016 annual report of the Australian Joint Registry, namely, that concern is raised about a type of prosthesis sufficient to justify further investigation, if the risk of revision was recorded as twice the average. Attractive though that proposition may appear, it would be dangerous for the Court to attempt to set any parameters for materiality in circumstances in which the claimants have deliberately avoided doing so. In any event, for reasons that will become apparent, it is unnecessary for me to embark on that exercise.

**THE RELIABILITY OF CRR AS A MEASURE OF SURVIVORSHIP**

280. Mr Whitwell and Professor Pandit agreed that an adequately powered randomised controlled trial and/or a systematic review of such randomised controlled trials with sufficient follow-up would be the “gold standard” for comparing survivorship as well as other outcomes between different types of implants. Ms Smith also agreed that a large scale randomised control study, if sufficiently well designed, would be the best source of data for comparing the behaviour of different prostheses. Professor MacGregor said that although a randomised control trial of prostheses would be an optimal study design, realistically, it was not going to be feasible to conduct such a trial and carry out follow-up to ten years. Therefore, one must necessarily rely on data that provides that information from other sources.
281. It was accepted by all relevant experts that observational studies, based on data entered in the registries, will be less reliable than a randomised control study. The generic orthopaedic experts agreed that CRR can enable a comparison of survivorship to be made between different types of implants when the data are complete and accurate, though they are not the only way of assessing implant performance.
282. A critical aim of a statistical study investigating a covariate in two populations is to try, so far as possible, to ensure that the populations are identical, save for the covariate being studied. Therefore, it would be important to know whether there were any factors that might bias or confound the comparison, and thus adversely affect its reliability. When looking at the data, one must take into account and make appropriate adjustments, *where possible*, for bias and potential confounding factors. I stress “where possible” because it may be impossible to evaluate how the potential confounding factors would impact upon the collated data, for example, because there is no data relevant to those confounding factors. Ms Smith readily acknowledged that direct comparisons between different types of hip prosthesis were not straightforward because of differences in the characteristics of the patients in each group. Professor MacGregor agreed that any data obtained from registries had to be analysed carefully with an eye on the biases and confounders that may affect that data.
283. As Ms Smith explained in her first expert report, in order to measure the risk of failure or the need for revision over time, it is necessary to carry out a “survivorship analysis” which requires accurate data to be recorded on both the primary and revision operations. Those two sets of data must be linkable (to ensure that it relates to the same patient and the same prosthesis), and a reasonable amount of time must pass since the primary surgery in which all patients are observed. Links must also be made with mortality data so that patients stop being treated as being at risk of revision (the identified “hazard”) once they have died. Survivorship analysis can be presented in different ways – as revision rates at particular points in time; as patient time incidence rates; or as cumulative hazard (risk) graphs.
284. It is important to bear in mind that what the statistical evidence seeks to do is evaluate the cumulative risk of a hazard occurring over a given period of time, in this case 10 years, by calculating the probability that patients implanted with a particular prosthesis would have it revised within that period. Although the prediction is based on data relating to the actual revisions of primary implants each year, it will not be possible to ascertain how accurate the prediction has proved to be in practice until many years of data have accumulated. As the calculation will include a prediction of

the percentage of implants that will need revision within 10 years from a primary operation taking place in year 10, it will only be possible to establish what percentage of the hips implanted in years 1-10 in fact required revision within 10 years, when there is 20 years' worth of data.

285. There is likely to be some degree of error or inaccuracy in any data set. What is important is the extent of any error. A confidence interval is the way in which statisticians calculate the margin of error (plus or minus) around the value of what is being measured, in this case, the CRR. One way to reflect how certain it is that a sample estimate (or "point estimate") reflects that value is by using a confidence interval for a specified confidence level. The confidence level chosen, and the standard deviation of the sample estimate, determine the calculation of the margin of error; typically, the wider the confidence level, the wider the margin of error. The size of the population from which the data is taken will also affect the margin of error; a larger sample set will reduce the margin of error and vice versa. Confidence intervals are usually referred to by their chosen value of confidence level, e.g. a 95% confidence interval is one with a confidence level of 95%. If 95% confidence intervals are calculated on each occasion the data is analysed, then 95% of those intervals will contain the true value of what is being measured. In this case, 95% confidence intervals were used.
286. It is crucially important that any method used properly takes account of the amount of time that a patient has been observed and is at risk of revision. This requires the use of advanced statistical techniques to take account of censoring (where patients are not all observed for the same length of time). Ms Smith chiefly relied on the Kaplan Meier survival curve, which estimates the probability of surviving at a given point in time. The analysis aims to estimate a survival curve in a given population from the sample of patients in the presence of censoring. Although that type of analysis is commonly used in joint registries, it is limited because it does not adjust for covariates (other factors that can influence results). If such adjustment is needed, more complex multivariable statistical models must be used instead.
287. A Kaplan Meier analysis calculates the risk of a hazard, in this case, the risk of revision, for each individual patient for each day from the date of primary surgery (which is when the patient becomes at risk of revision) until the patient is revised or "censored", that is, ceases to be monitored for any reason (e.g. they die or they are lost to follow up), until the end of the observation period. At each time the hazard is calculated as the number of patients who were revised over the number of patients at risk. The analysis assumes that the risk of revision among those who are censored at a given time is the same risk faced by those who remain at risk at the time the hazard is calculated. To calculate the cumulative rate, the statistician adds all the individual risks together. This means that every single patient contributes to the data, even if they have only been observed for one day.
288. The analysis will be subject to change as more data become available over time. Thus, the calculation of the 10-year cumulative survivorship/revision rates of specific hip articulations between 2002 and 2012 as at December 2014 may produce different figures from the exercise carried out in 2012, or in 2017 or 2021. However, as Ms Smith explained, in assessment of the reliability of the risk of revision it is the confidence intervals rather than point estimates that matter. If the confidence intervals

overlap, then there is no statistically significant difference between the rates reported: the differences are likely to be due to random variation.

289. The first step for the Court to take is to ascertain what the appropriate comparator product or products should be, for the purposes of carrying out the comparison of their CRR. The second is to identify the relevant CRR for the comparator(s) and the Pinnacle Ultamet, assess their reliability, and carry out the comparison exercise. The third is to consider whether it can be reliably concluded on the evidence that there was a materially increased risk of early revision for the Pinnacle MOM implant. If the answer to that question is yes, the final step is to consider whether in the light of that materially increased risk and all other relevant circumstances, the product fell below the level of safety that the public generally was entitled to expect of a prosthesis used for a total hip arthroplasty at the time when it entered the market in 2002.

## **4.2 THE APPROPRIATE COMPARATOR**

290. In terms of appropriate comparators by which to evaluate the CRR, the claimants sought to rely upon:
- i) A comparison between the Ultamet and the other forms of liner within the Pinnacle modular system itself (“the intra-Pinnacle comparison”) and/or
  - ii) A comparison between the Pinnacle Ultamet and prostheses with alternative bearing surfaces (“the external comparison”).
291. The Directive and the Act make it clear that a defect in a product cannot be inferred from the fact alone that a product which enters the market subsequently has better or enhanced safety features. Where the complaint is of an abnormal risk of early failure, compared with a prosthesis that the patient might have received instead, and especially bearing in mind that at the time the Pinnacle Ultamet was introduced all producers were aiming to improve on the durability of the existing products, the performance of later products cannot inform the assessment of the entitled expectation of safety. The current generation of CoC products were introduced some years after the Ultamet. Some of the new generation of cross-linked and HXLPE alternatives to conventional polyethylene were coming onto the market at around the same time as the Ultamet, but their scientific development continued throughout the 2000s. Different forms of HXLPE undergo different levels of intentional cross-linking and different stabilisation techniques, and so different versions of HXLPE were produced by different producers at different times. Therefore, it would be difficult to find a suitable comparator or comparators from that class. Moreover, the pattern of usage by surgeons of the different articulations shifted over time. For example, Mr Whitwell himself did not start using HXLPE implants until 2010.
292. Some of the products whose performance is included in the statistics upon which the claimants relied, were introduced later. Indeed, some even came onto the market after the Ultamet liner was withdrawn. When looking at the statistics, therefore, caution is necessary if and to the extent that data relating to MoP articulations includes data relating to the newer types of cross-linked polyethylene or HXLPE or newer products. Unfortunately, the statistics do not inform us about the extent to which the data in the

NJR include data on such products, save for the intra-Pinnacle data supplied by DePuy.

### **THE INTRA-PINNACLE COMPARISON**

293. In my judgment the intra-Pinnacle comparison is inapposite. I understand why the claimants wish to rely upon it; by doing so, they would eliminate any potential confounders that might be introduced by differences in product design, or differences between cemented hips, like the Charnley, and uncemented hips. However, that is no justification for using an otherwise inappropriate comparator. The public were not entitled to expect any one of a group of new products introduced onto the market at about the same time as the Ultamet to be safer (or “materially” safer) than any of the others, though they would be entitled to expect all of them to achieve at least a threshold level of safety. In the absence of something that would entitle anyone to expect that the new product would have an elevated level of safety (for example marketing claims by the manufacturer) that threshold level is likely to be measured by reference to the minimum level of safety that was then expected of existing comparable products.
294. The entitled expectation of safety must be evaluated by reference to what was known or might be expected in terms of safety at the time the product went onto the market, including any information in the marketing materials or packaging of the product that warned of potential risks or hazards. A comparison with the subsequent performance of other new products – in this case, the other articulations within the Pinnacle system - could not inform the entitled expectation. To use information gathered since 2003 about the other Pinnacle articulations for that purpose, is an inappropriate use of hindsight.
295. Using another new product as the comparator would also lead to the absurd conclusion that even if all the new products showed an improvement on the existing established products in terms of safety, the new product that showed the smallest improvement by comparison with the others could nevertheless be regarded as defective, if the difference between them was of a sufficient magnitude. Mr Whitwell agreed that, if this approach were to be taken, a manufacturer with more than one product on the market might be wise not to market a new product which might perform particularly well, given that others in the range might not compare favourably with it. That would be anathema to research and development.
296. Mr Oppenheim submitted that that was not this case, and if hypothetically there were a case in which the worst performing product in the same modular range produced results that were better in terms of survivorship than external comparators, the Court might then stand back and decide that that product was not defective. That argument does not engage with the fundamental flaws in the use of the intra-modular comparison. If the Court must look at the external comparators anyway, there is nothing to be achieved by carrying out an intra-modular comparison. If the new product performs materially worse in terms of safety than an appropriate external comparator, there is no necessity to look within the modular system to assess whether it falls below the entitled expectation of safety. If it performs as well as or better than the external comparator, which *ex hypothesi* already meets the entitled expectation of safety, the new product cannot fall below the entitled expectation of safety merely

because it performs materially worse than other articulations within the same modular system.

297. Mr Oppenheim made the point that MoP hips were not going to be withdrawn from the market and that the new and old products would co-exist. That is no reason to confine the comparison to other articulations within the same modular system; if anything, it is a good reason to compare the new products with the existing ones. If safety is to be measured by the risk of early failure, the public would be entitled to expect that the new implants would not carry with them a risk of early failure that was materially worse than that expected of the tried and trusted products that were already on the market. The appropriate comparator should therefore be a product or product type that was already on the market, and which had been on the market for long enough to engender sufficient robust and reliable data about revision rates and survivorship to inform how long a person might reasonably expect their hip prosthesis to last before requiring revision, irrespective of who the manufacturer was.
298. There is also force in Mr Antelme's point that the intra-Pinnacle comparison could only be made because DePuy chose to offer non-MoM alternative bearings – if a producer simply made MoM implants, it would not be possible. That raises the unedifying prospect of other defendants to similar claims being subjected to an analysis of safety by reference to a different comparator, which would be neither fair nor consistent. Insofar as the Pinnacle system included the option of a metal on conventional polyethylene articulation, the data relating to that articulation will be included in the body of data concerning MoP prostheses generally.
299. Mr Oppenheim strongly relied on the fact that the Pinnacle system was specifically marketed as a modular system with the flexibility for the surgeon of being able to have a wide choice of different articulations. Whilst that is true, it does not justify use of the alternative articulations within the modular system as comparators for the evaluation of the entitled expectation of safety. The public would not be entitled to measure their expectation of safety merely by reference to other products supplied by the same manufacturer. DePuy did not have a monopoly, and older products were still available. I therefore reject the suggestion that the appropriate comparison is with all the other articulations that were available within the Pinnacle modular system.

### **THE APPROPRIATE EXTERNAL COMPARATOR**

300. The Pinnacle modular system was an uncemented prosthesis. If one is assessing what the public would be entitled to expect in terms of safety of an uncemented prosthesis with a large head MoM articulation coming onto the market in 2002, the obvious comparator would be an uncemented prosthesis. The aim is to compare like with like, so far as possible. Logically, in order to eliminate any differences caused by fixation method, the comparator should be a prosthesis with the same method of fixation.
301. A further reason for reaching this conclusion is the demographic of the group who received the large head Pinnacle Ultamet implants. DePuy submitted that the best comparison will be a comparison with an established uncemented metal on conventional polyethylene prosthesis, as that would be the type of prosthesis which would have been provided to younger and active patients before the new generation of alternative bearings became available. The Pinnacle Ultamet was developed



specifically with younger and more active patients in mind, with a view to improving survivorship beyond 10 years: they were the target cohort.

302. At the time when the Pinnacle Ultamet was introduced, and, indeed, thereafter, most elderly and inactive patients in the UK would have been given a Charnley hip or similar cemented prosthesis. The data from the NJR indicates that even after 2002 when both types of prosthesis were available, an older patient was far more likely to have been given a cemented MoP hip in preference to a Pinnacle Ultamet, unless, of course, there were specific reasons for choosing a large head MoM articulation (such as a desire for increased stability). The choice of articulation often falls to the surgeon, who will generally have his own preferences. Mr Whitwell, for example, continued to use metal on conventional polyethylene articulations for most of his patients until around 2010, and still does for his older patients.
303. In 2002 a total hip replacement was unlikely to have been offered to someone aged under 70 *at all* unless the indications for an arthroplasty left the consultant with little choice. The younger patient might well have been offered a resurfacing arthroplasty instead of a MoP hip. The younger and more active patients who would have had a total hip arthroplasty both before and after 2002 would have been more likely to receive an uncemented prosthesis than a cemented one. That was also true of the patients in Sweden whose data contributed to the statistics recorded by SHAR.
304. The received wisdom as to the anticipated survivorship of a hip prosthesis at the time was based on the track record and predicted future track record of all MoP hips and the CRR as recorded in the SHAR 2000 report. The overwhelming majority of these prostheses (around 93%) were cemented. Data about the performance of cemented prostheses, or which largely comprises data concerning their performance, must be treated with some caution, not just because there is a different method of fixation but, more importantly, the patient demographic receiving such implants is largely unrepresentative of the cohort which would be receiving Pinnacle Ultamet prostheses. On a broad overview of all the statistics in this case, cemented implants generally outperformed uncemented implants in the first 10 years after primary surgery, irrespective of the patient's age. The 10-year CRR for cemented prostheses with a MoP articulation in all patients has always been in the region of 5% or lower.
305. Mr Oppenheim rightly pointed out that the IFUs for the Pinnacle Ultamet were not restricted to younger patients. He contended that, at least up to the early 2000s, younger patients were more likely to have revisions because of more complex underlying diagnoses than osteoarthritis, which is the most common underlying diagnosis in older patients. He said that over the entire 9-year period that the Pinnacle Ultamet was on the market, over 88% were implanted in patients aged over 55, and the average age of patients with such implants over that period was 67. Many Pinnacle Ultamet prostheses were implanted in patients over 70. However, Mr Oppenheim's 88% included all the patients who were aged 55-69, some of whom would have been more active than others.
306. Those figures need to be put in context. Mr Spencer QC showed Ms Smith a table produced by Professor Scharfstein showing the distribution of Pinnacle Ultamet bearings by age and sex to three end points. She said she had no reason to quarrel with the data in that table. She accepted that there was an upward trend in the average age of patients offered the Pinnacle Ultamet prosthesis over time, and that the mean and

median ages of such patients in the earlier years were younger than when measured over the entire period of implementation. In the period up to August 2005 the median age was 62.4 years; a year later it went up to 64, culminating in 66.5 in 2011. Whatever the average ages of the Ultamet cohort may have been at different times, most were aged under 70.

307. Mr Oppenheim submitted that because the Pinnacle Ultamet prosthesis was implanted across the entire spectrum of patients, the comparison should not be confined to data in respect of uncemented MoP hips, or data in respect of younger patients. However, the fact remains that there will have been a much larger proportion of patients aged under 70 and more active patients in the group who received the Pinnacle Ultamet, especially in the early years, and they are the ones whose activity levels might be expected to (and were expected to) raise the revision rates. That is one reason why the overall statistics for all cohorts of patients with MoP implants prior to 2002 will not be a reliable direct comparator, because the overwhelming majority of the patients contributing to that data were over 70. The overall CRR are therefore likely to be a significant underestimate of the CRR that would have been produced if the Ultamet patients had been implanted with an alternative prosthesis.
308. Given that the Pinnacle system was itself an uncemented implant and that uncemented prostheses were likely to have been used for those patients who would have made up the larger proportion of the cohort implanted with a Pinnacle Ultamet prosthesis, I consider that the figures for uncemented prostheses are the best available and most appropriate to use for comparative purposes, though they are far from perfect. This conclusion is supported by Professor MacGregor's ready acceptance in cross-examination that from an epidemiological point of view it would be strange to use a cemented product as a comparator for the Ultamet. However, in view of the fact that the Pinnacle Ultamet prosthesis was implanted across all age groups, the material should not be confined to data relating to patients aged under 70. The Court must also bear in mind that the data on cemented MoP prostheses would improve the figures for survivorship, to the extent that such data reflects the likely performance of implants that would have been received by the elderly patients within the Ultamet cohort.
309. At the time when the Pinnacle Ultamet was introduced, a patient undergoing a total hip arthroplasty was more likely to be offered a metal on conventional polyethylene articulation than any other type. Mr Whitwell said that there were cost implications with CoC articulations, which were significantly more expensive. There was also the risk of fracture, which had been reduced but not eliminated, and which made some clinicians (including Mr Whitwell) chary of recommending them. Ceramic on conventional polyethylene was the least popular combination at that time.
310. MoP was also the articulation whose known track record generally informed expectations in terms of survivorship at the time, and whose performance in terms of survivorship (as well as other features) the designers of the newer types of articulation were aiming to improve upon. That much was also clear from the literature and marketing materials for the Ultamet, which spoke of the potential benefits of the Pinnacle Ultamet "*compared with conventional metal-on-polyethylene implants*".
311. Realistically, when discussing the various options with a patient at that time, it is most likely that the orthopaedic consultant would be explaining the potential benefits of a MoM hip in comparison with a MoP hip. A MoP hip also contained the highest

wearing liner material of any of the articulations, which was also liable to degrade within the body over time, and therefore was likely to have the worst track record in terms of survivorship, but the claimants necessarily accept that its performance met the entitled expectation of safety. Therefore, the obvious comparator is an uncemented prosthesis with a MoP articulation using conventional polyethylene.

312. I do not regard other types of articulation as appropriate or necessary comparators if safety is to be measured in terms of short-term survivorship. Admittedly some of the group of patients who received Pinnacle Ultamet prostheses may have received a CoC or CoP implant instead, but it is impossible to tell how many. Some of them may have had a resurfacing operation instead. An equally unknown number would not have had a total hip arthroplasty at all. Any hip prosthesis containing third-generation ceramic material carried with it a higher risk of component fracture but a lower risk of aseptic loosening than MoP, which introduces unnecessary complications. Judging by the NJR data, the incidence of use of the alternative articulations using ceramic material was far smaller than for MoP. There was no long-term data relating to their survivorship over 10 years available in 2002; later data will be complicated by the introduction of the fourth-generation ceramic materials. In any event, the NJR data relied upon by the claimants indicates that the CRR of such other articulations were better than, though not significantly different from, the CRR they rely upon in respect of the performance of MoP, so looking at those other articulations is not really going to take matters further.
313. If data relating to MoP hip prostheses includes both metal on conventional polyethylene and metal on cross-linked polyethylene or HXLPE articulations, as it undoubtedly does in the NJR database and in the SHAR 2014 report, it is likely to be less reliable for comparison purposes than if it relates to metal on conventional polyethylene alone. The whole purpose of the intentional cross-linking was to produce a harder wearing and more durable surface; if and to the extent that succeeded (as the short-term scientific information on its performance suggests that it has) one would expect data relating to the cross-linked product or products to improve the CRR even in the short-term. There was some debate at trial about whether the inclusion of HXLPE in the NJR data made a statistically appreciable difference to the CRR for MoP over 10 years. Bearing in mind all the non-statistical evidence, as well as the history leading to the development of the alternative harder bearing surfaces, one would expect the data to demonstrate metal on HXLPE articulations to have an appreciably better CRR than metal on conventional polyethylene.
314. Despite this, in her supplemental report, Ms Smith had carried out a calculation based solely upon data supplied for the Pinnacle cup with a Corail stem, which was broken down between metal on conventional polyethylene and metal on HXLPE articulations, from which she concluded that over 9 years the CRR for conventional polyethylene were 2.65% and for HXLPE 2.23% and that there was no statistically significant difference between those figures. When asked about this in cross-examination, she said she would expect the CRR to be better for HXLPE, over 20 years, but not necessarily in the short term, and she was not surprised by these results. It then transpired that she mistakenly thought that the material in conventional polyethylene may have improved since the data reported in the SHAR 2000 report. She did not agree that the similarities in the CRR could be explained by HXLPE being given to the more active patients and conventional polyethylene to less active

patients (though other evidence suggested that this is precisely what would have happened; indeed, that was Mr Whitwell's own practice from 2010 onwards).

315. The relatively high CRR both before and after 10 years in uncemented hips implanted in younger patients, identified in the SHAR reports for 2000 and 2002, were undoubtedly ascribed to osteolysis by the orthopaedic community. Osteolysis was not a problem that only developed in the longer term. I have already referred to Dr Haas's evidence about his own experience and that of his surgical colleagues regarding the development of osteolysis, with which the generic orthopaedic experts broadly concurred. Typically, it would arise at around 7-10 years after the primary implant, but it could arise earlier than that if the implant were subjected to greater wear. If there was insufficiently regular follow up with conventional X-rays, the condition would only be detected at a later stage, but it could have arisen much earlier. The development of harder wearing alternative articulating surfaces, including cross-linked polyethylene and HXLPE, was driven by the incidence of osteolysis, although the focus was on longer-term survivorship.
316. Ms Smith's conclusions that there is no appreciable difference between the predicted 10-year survivorship of conventional polyethylene and HXLPE do not fit with all the other evidence that the Court heard about HXLPE and the fact that other scientific studies, such as the *in vivo* study carried out in Oxford by Dr Glyn Jones and his team (all of which were necessarily short-term) suggested that at least one type of HXLPE was performing much better than conventional MoP. If Ms Smith's conclusions based on her analysis of the intra-Pinnacle data in the NJR were a reliable indicator of how HXLPE compared with conventional polyethylene, then as Mr Antelme pointed out, either the whole history of prosthetic development in the 1990s and 2000s was a complete waste of time, or else DePuy fortuitously happened to develop a particularly hard-wearing brand of conventional polyethylene (which was not the case).
317. This illustrates just how careful the Court needs to be with drawing inferences from CRR, particularly when dealing with relatively small amounts of data relating to articulating surfaces within one modular prosthesis over a 9-year period. Ms Smith was generally very careful not to trespass beyond the bounds of her expertise. The figures on which she relied were predictions of the likelihood of failure within 9 years of prostheses using different articulating surfaces. She could not give evidence about how these materials behaved in the real world, as the orthopaedic surgeons could and did. Bearing in mind their evidence, I cannot accept that HXLPE made no appreciable difference to the survival rates of MoP articulations over the first 10 years after primary surgery. It is far more likely that it improved them, and thus that the inclusion of significant amounts of data pertaining to articulations using HXLPE would have materially improved the 10-year CRR for MoP articulations, to an unknown and unascertainable extent.

#### **THE CRR OF THE EXTERNAL COMPARATOR**

318. It is not easy to find data which enables an evaluation to be made of the likely survivorship of an uncemented MoP hip to 10 years or more following implantation (or the probabilities of revision within that period) in those patients who would have been given an uncemented metal on conventional polyethylene hip, if they were given a prosthesis at all, but who were instead given a Pinnacle Ultamet in the 2000s ("the Ultamet cohort"). In an ideal world, one would be looking for data on revisions or

survivorship of an uncemented metal on conventional polyethylene prosthesis or prostheses implanted in a group of patients which would include a sufficient proportion of younger and more active patients to be representative of the cohort that was implanted with a Pinnacle Ultamet prosthesis. Unfortunately, that data does not exist. That is not the claimants' fault, but it does create a major problem for their case, because it makes the crucial comparison exercise extremely difficult.

319. As Ms Smith and Professor MacGregor stated in the joint expert statement, there is relatively little large-scale, representative data available on the survivorship of metal on conventional polyethylene bearings in 2000 and for the period 2001-2009. They accepted that data from the Nordic registries such as SHAR is probably the most appropriate for those purposes, although there are differences in population characteristics, surgical techniques and fixation between the Scandinavian countries and the UK which make direct generalisation difficult. Neither Ms Smith nor Professor MacGregor carried out any analysis of the performance of the established MoP implants on the market at the time when the Pinnacle Ultamet was introduced, for the purposes of the comparative exercise, but Ms Smith was asked about the SHAR data in cross-examination.
320. Prior to the introduction of the new generation of hip prostheses, if a younger patient did have a total hip arthroplasty they would have been warned not to return to high activity levels. Thus, even historical data for a younger cohort receiving uncemented metal on conventional polyethylene prostheses would not completely reflect the likely impact of the activity to which a Pinnacle Ultamet prosthesis was likely to be subjected – as illustrated by some of the test cases, in which patients returned to regularly playing sports, or walking long distances.

#### **4.3 THE SWEDISH HIP ARTHROPLASTY DATA**

321. Although the SHAR data were the only long-term data available at the time when the Pinnacle Ultamet was introduced, and the data in the 2000 SHAR report informed the views of orthopaedic surgeons and NICE as to the likely survivorship of MoP prostheses, Mr Oppenheim submitted that the 2000 and 2002 SHAR reports provided insufficiently reliable data for the purposes of making the comparison. I do not accept that submission, which flies in the face of the evidence of his clients' own expert statistician and epidemiologist.
322. Whatever its limitations, SHAR was the registry which informed the orthopaedic community's understanding of survivorship of hip implants at the relevant time. Both generic orthopaedic experts confirmed its importance as a source. Mr Whitwell said in his evidence that "*we still quote from it quite reliably because it's the longest registry*". I find that the SHAR registry is the most appropriate source of relevant and reliable information about the actual and likely performance of conventional MoP prostheses over the short and longer term. The NJR, which was only set up in 2003, has no data pertaining to any 10-year period up to and including 2002, when the Ultamet came onto the market, and its data relating to the performance of MoP prostheses over later periods is confounded to an unknown and unascertainable extent by the inclusion of data pertaining to HXLPE and products that came onto the market later than the Pinnacle Ultamet, even before one even begins to consider other confounding factors. If the SHAR registry is an unreliable source of information

about the 10-year survivorship or CRR of MoP prostheses for the purposes of the comparative exercise, the NJR is even less reliable.

323. In cross-examination, Ms Smith agreed that the data in SHAR 2000 for the uncemented prostheses in younger patients were a reasonable comparator for the Pinnacle MoM cup, but she said that because uncemented prostheses were far less common in Sweden than cemented prostheses, she was concerned that there might be something atypical about the uncemented implant group in Sweden. That concern can be tested by reference to the data in the 2002 SHAR report for the uncemented implant group, which strongly indicates that any atypicality relates specifically to those prostheses that failed for reasons other than aseptic loosening in patients with diagnoses other than osteoarthritis.
324. Ms Smith did not say that the cohort of patients implanted with uncemented prostheses in Sweden was statistically too small to provide reliable data. She simply pointed out that it was much smaller than the cohort implanted with cemented prostheses in Sweden. That explains why the confidence intervals are wider than they are for the cemented prostheses, but they are generally comparable with the range of confidence intervals in the data in the Australian National Joint Registry, which is accepted by Ms Smith to be a useful secondary source.

#### **THE SHAR 2000 REPORT**

325. The SHAR 2000 report is not a perfect source, not least because it does not count as a “revision” the exchange of the femoral head or liner in a modular system. Moreover, it was confined to patients whose primary presentation was for osteoarthritis, thus excluding other patients who may present greater surgical challenges and a correspondingly higher risk of revision. Mr Whitwell accepted this. He also accepted that because of this, the recorded CRR were minimum revision levels, because other information in the report indicated that there was an increase of 1-2% for other causes of revision besides aseptic loosening in the cemented hips. Although there were no comparable figures given for other causes of revision of uncemented hips, he agreed it was a fair assumption that the increase would be of a similar magnitude. I am satisfied that this is a reasonable assumption; indeed, there is evidence in later SHAR reports to support it.
326. Whilst the tables in the SHAR 2000 report cannot provide a direct and entirely reliable comparator with the Pinnacle Ultamet, they do provide an indication of the predicted rates of revision/survivorship of existing MoP prostheses at 10 years and 16 years, and of the dramatic increase in the CRR after 10 years that had a considerable influence on the orthopaedic community before and at the time when it was introduced.
327. Ms Smith said in her initial expert report that the Swedish registry had data on around 400,000 hip arthroplasties and could provide long-term survivorship analysis. The sole reason she eliminated SHAR as a reliable basis for carrying out the full comparative exercise was that data specifically relating to the Pinnacle cup, in the latest SHAR report for 2014, was not split by bearing surface. However, the fact that SHAR may have insufficiently reliable data to enable a comparative assessment of the short-term survivorship of the various articulations within the Pinnacle prosthesis does not rule it out in terms of providing reliable data relating to the performance of

*external* comparator prostheses that were on the market when the Ultamet was introduced.

328. The average age of patients in the Swedish registry at this time (70) was older than the average age of those implanted with the Pinnacle Ultamet at any time. The Swedish age profile did, however, correspond with the general demographic of patients who would have received total hip arthroplasties in the UK before the Ultamet was introduced. It will be seen that the data for the younger and more active patients in Sweden who formed the minority of patients in that country had an insignificant statistical impact on the overall CRR – it was only when these data were segregated that it became clear that the overall revision rates were wholly unrepresentative of the revision rates that could be predicted for the younger patient cohorts. The same is likely to be true if the demographics of the patient group were reversed. Thus, although the higher actual and predicted revision rates for the younger and more active patients who would have formed the majority of the Ultamet cohort would be balanced to a limited extent by the lower revision rates for the elderly and less active patients who formed the minority, it is unlikely that the latter would have had a significant impact on the CRR for the entire cohort.
329. The SHAR 2000 report showed that, using a Kaplan Meier analysis, the predicted survival rate over 10 years across the board for uncemented implants (though only in patients with osteoarthritis who were revised for aseptic loosening) for the period from 1988-1998 was a point estimate of 85.8% with confidence intervals of 82.9% to 88.8% (a CRR of 14.2% with a margin for error of around 3% either way). Cemented implants, over the same ten-year period, had far better predicted survival rates of 94.6% with confidence intervals of 94.2% to 95.0% (a CRR of 5.4%, with much smaller margins of error). This is illustrated by the graphs in the report. As I have already mentioned, for all hip implants the analysis indicated a sharp decline in survivorship after 10 years. Ms Smith agreed that, whilst it looked from the graphs as if the CRR for these implants might be improving in the 1992 cohort, after the uncemented surgical technique had changed, there was only data available to about six years at that time.
330. The SHAR 2000 report also indicated that the predicted rates for survivorship of existing implants in younger patient cohorts were worse than the rates overall. The report stated that approximately 12,000 young patients were included in the entire cohort (1979-1998) with a high failure rate, and that it was for these patients that further scientific effort was mandatory. In the under-55 age group the graph for revision for all causes showed a survival rate at 10 years of 81.2% with a confidence interval of 78.4-84.2% for implants in male patients, but after 16 years that had plummeted to 32.9% with a confidence interval of 31.5%-34.3%. The statistics for female patients were slightly worse.
331. There is a degree of correlation between the statistics for the younger age groups and the data for the uncemented hip implants, which is unsurprising, since they were the patients who were most likely to have received such implants. Ms Smith accepted that in the period from 1988-1998 the risk of revision within ten years for the younger groups was calculated to be around 20% for both sexes. That was the data which led the authors of the SHAR 2000 report to conclude that younger and more active patients were at greater risk of implant failure in all diagnostic groups. That was also the received wisdom in the orthopaedic community, supported by practical

experience, at the time when the Pinnacle Ultamet was introduced onto the UK market.

332. Mr Oppenheim submitted that the higher risk of failure in younger patients was not necessarily due to increased activity and higher wear but could have been due to the poor fixation techniques in Sweden, and there was no way of knowing how much effect the poor fixation techniques had on the data. However, the poor fixation techniques should have affected all patients who were implanted, irrespective of age; moreover, they appear to have been more of a problem with the cemented implants. The narrative in the report indicated that uncemented technology had a disappointing result in the cohort operated on prior to 1988, but that results had improved since then, especially in the post-1992 period when modern cup designs and active surface coating on the femoral component were used. It stated that “*the third generation of uncemented implants used in the nineties has functioned relatively well with excellent fixation*”. There is no reason for me to doubt those statements.
333. In any event, consistently with the practical experience of the orthopaedic community in the preceding decade, the authors of the SHAR 2000 report identified wear as the main cause for concern, driving the need to find alternative articulations for younger patients.
334. The generic orthopaedic experts understandably focused on the data for the period from 1988-1998 in their reports. The fixation techniques would not have had a significant impact on the recorded CRR post-1988, at least for the cemented hips. In order to test whether the fixation of uncemented prostheses unduly distorted those statistics, the predicted survivorship for hips implanted during that period should be compared with the figures for the 10-year period from the time when those fixation techniques were improved, i.e. 1992. These are found in the SHAR 2002 report.

### **THE SHAR 2002 REPORT**

335. The next SHAR report (“the SHAR 2002 report”) was published in April 2003. In that report, which was compiled with the benefit of the additional data collated since 1998, the definition of “revision” now included changes in prosthetic components. Just as in the 2000 SHAR report, wear, rather than poor fixation, was identified as the main cause for concern. By then there were 10 years of data available since the introduction of the changes to uncemented fixation techniques in 1992. The 2002 report states that developmental work on solving the problem of wear, with subsequent osteolysis and prosthetic loosening in younger patients, must continue, with the focus on alternative bearing surfaces that are more resistant to wear. Like the 2000 report, this is consistent with all the other evidence about the concerns that had been driving the search for alternative articulations during the preceding decade.
336. The report contains *overall* cumulative predicted rates of survivorship, this time for all implants, all diagnoses and all reasons, of 92.8% at 10 years from 1992-2002 (confidence intervals 92.4% - 93.1%), equal to a CRR point estimate of 7.2% and a range of 6.9% -7.6%. In other words, the analysis of the data indicated a statistical probability that between 92.4% and 93.1% of patients implanted with a prosthesis at any time within that 10-year period would not require revision surgery for at least 10 years. Conversely, there was a risk that between 6.9% and 7.6% of those hips would be revised within 10 years. That was the most recent period for which 10-year



statistics were available at that time. The graphs for that period indicated a slight improvement over the 10-year figures for the survival curves of implants where the primary operation took place in 1986 and 1989 (which are very similar to each other), a greater degree of improvement over the survival curve for primary implants in 1983, and a vast improvement over the survival curve for primary implants in 1979. Apart from the 1979 cohort, the predicted survivorship of all cohorts to around 7 years is comparable, and then the later implants begin to show an improvement.

337. As with the 2000 report, over 90% of MoP implants contributing data to those overall statistics would have been cemented implants used in elderly inactive patients. When the statistics are broken down, it can be demonstrated how unrepresentative the overall percentages are of the performance of an uncemented MoP implant at that time, irrespective of the age of the patient. For cemented implants, the extrapolated figures for the 1992-2002 period were 94.0% over 10 years (93.7%-94.4%), a little better than the overall figures for all implants referred to above, but for uncemented implants the same figures at 10 years were 77.6% (74.6-90.7). That is the equivalent of a CRR of 22.4%, despite the change in surgical techniques in 1992, though the top end of the confidence intervals suggested that the CRR could be much lower, 9.3%. These rates are much worse than the six-year CRR for patients with uncemented implants in the 2000 SHAR report. The confidence intervals are also far wider than with other cohorts, possibly indicating that the data is based on small numbers. However, the figures for all uncemented implants and all causes of revision are also far worse than the figures for revisions for aseptic loosening in patients with osteoarthritis. From this it can be deduced that the poor overall CRR are linked with revisions for fractures, infection or dislocation. The overall figures for uncemented implants in the SHAR 2002 report therefore need to be treated with circumspection.
338. None of these figures are directly comparable with those in the 2000 SHAR report. The figures which *are* comparable (though not exactly, as the later data include exchanges of components in the definition of revision) are those for revisions of implants for aseptic loosening at 10 years for patients with a primary diagnosis of osteoarthritis. For uncemented implants, the survival rates in the 2002 SHAR report were 83.3% (80.1%-86.6%). Those figures are similar to the survival rates for uncemented implants in osteoarthritic patients revised for aseptic loosening over the 10 years from 1988-1998 recorded in the 2000 report. There is also an overlap between the confidence intervals, making the recorded CRR statistically consistent. I regard this as being of considerable importance. Once the unknown factor that affected the statistics for revisions for other diagnoses and causes is eliminated, the data reliably indicates that the CRR over 10 years for uncemented implants were much the same before and after the changes in surgical fixation introduced in 1992. Uncemented implants were still performing worse than cemented implants over the first 10 years and even on a conservative basis it was predicted that around 15% of them would fail in that time. This puts the CRR of 7.2% for all implants, 93% of which were cemented, into context.
339. Even if one sets to one side the overall figures for uncemented implants on the basis that they appear to have been skewed by some reason or reasons for revision other than aseptic loosening, it is still a fair assumption that, based on the SHAR reports to April 2003, the CRR for uncemented implants over the most recent 10-year period were still, *at best*, somewhere in the region of 15%, notwithstanding the change in

surgical techniques for uncemented hips in 1992. For cemented implants, by contrast, the CRR was around 5% and for all types of prosthesis it was in the region of 7%.

340. I note in passing that if one applied the claimants' approach to the comparable CRR over the first 10 years as being reliable measures of the failure rates of those implants, and thus as a reliable means of assessment of the relative safety of cemented and uncemented prostheses, the magnitude of the difference between them might be said to demonstrate that all uncemented implants failed to meet the entitled expectation of safety in 2002 because these statistics establish that they were three times more likely to fail within 10 years of implantation than a cemented implant. Even though it was recognized at the time that the rates of failure of all prostheses, but particularly the uncemented prostheses, needed to be improved upon for younger patients, the reason for this was the high incidence of wear *after* 10 years. The public were plainly not entitled to expect those prostheses to have a failure rate within the first 10 years that was "not materially higher" than that of the cemented ones. This illustrates how important it is to be cautious about the use to which one puts statistics.
341. So far as the age of the patients is concerned, there is no direct correlation between the tables of data in the 2000 and 2002 SHAR reports because the latter divides the age groups differently. For male patients aged between 50 and 59 years old at the date of primary surgery, the figures for 11 years are 85.2% overall (83%-87.5%), 84.6% (81.5%-87.6%) for cemented implants and 85.2 (80.5%-89.9%) for uncemented implants. The figures for female patients are generally slightly better, though for uncemented implants they are worse, 81.0% (74.4%-87.6%). Separate graphs for overall survivorship in the under-50s are worse than those for the group aged 50-59, namely 81.9% (78.6%-85.3%) for male patients, and 77.1% (73.0%-81.1%) for female patients. For uncemented prostheses, the figures for male patients are 77.9% (71.4%-84.4%) and the figures for female patients are 72.4 (66.0% - 78.8%). It is a fair assumption that those CRR would be higher in a cohort where many of the patients would have returned to higher activity levels than any Swedish patient, irrespective of age, would have been advised they could undertake before 2002.
342. The expert opinion of Ms Smith and Professor MacGregor in the joint statement was that in the SHAR 2002 report the revision rate for all implants was around 15% at 10 years, 20% at 15 years and 23% at 20 years post-operatively. Those figures include data from the periods prior to the changes in surgical techniques, but they are still consistent with the figures for uncemented implants across all age groups, and implants in younger patients, over the most recent 10-year period, 1992-2002, after the surgical techniques had changed.

### **THE 2002 MALCHAU PAPER**

343. Ms Smith was shown another paper written by Dr Malchau in 2002 in re-examination, though her evidence about it was limited. She did not suggest it cast any doubt on SHAR 2000 or SHAR 2002. This paper, like SHAR 2002, included exchange of a liner or head component in the definition of "revision". For cemented implants, in patients with osteoarthritis, survival rates for aseptic loosening over the intermediate 10-year period from 1990 to 2000 were reported to be 94.8% with confidence intervals of 94.4%-95.2%; but uncemented implants in such patients had a survival rate over the same period of 87.7% with confidence intervals of 85.2-90.3%. Over 17

years there was a similar “cliff edge” phenomenon to that observed in the 2000 and 2002 SHAR reports.

344. So far as the age of patients was concerned, the Malchau paper stated that over 9 years post-operatively, prostheses in patients with all diagnoses revised for all causes and aged between 55-75 years had a survival rate of 94.3% with confidence intervals of 93.8%-94.7%. For patients aged under 55 this dropped to 87.6% (85.9%-89.3%). However, this data only goes to 9 years and not 10 (the time at which the “cliff edge” effect starts to show on the graphs) and the age span in the first group is very wide - 20 years. There is no separate data pertaining to the group aged 60-69 which is likely to have had the widest mix of active and less active patients within it. The data from 1990-2000 also spans the period during which surgical techniques for uncemented implants changed in Sweden. Despite this, the Malchau 2002 paper does not appear to me to be inconsistent with the overall picture painted by SHAR 2000 or SHAR 2002 in respect of the predicted incidence of early revisions for aseptic loosening in patients with osteoarthritis. I need say nothing further about it.

### **THE SHAR 2014 REPORT**

345. SHAR did produce a further report in 2014, but the ten-year CRR in that report were based in part on data drawn from the new alternative bearing surfaces, including the different types of HXLPE. Moreover, the commentary in the body of the report specifically referred to the fact that the data for the most recent decade in the report (2005-2014 inclusive) included data relating to new products. The report indicates that HXLPE was used increasingly in Sweden from 2002 onwards. MoM was so rarely used that it is statistically insignificant. A table on page 34 demonstrates that metal on conventional polyethylene still accounted for just over 80% of primary implants in 2005, but its use then decreased over time. The use of metal on conventional polyethylene became roughly equivalent with metal on HXLPE in 2009-2010, after which HXLPE rapidly became more popular, gaining predominance in 2013. By 2014, HXLPE was used “almost exclusively” in uncemented cups and in just over 70% of cases of cemented cups.
346. The data also included patients of all ages and activity levels. The average age of patients had gone down in the period since 2000 and was now around 67 for men, 69 for women. Uncemented implants were growing in popularity, and still tended to be used in the younger patients, though cemented prostheses were still used in around 74% of primary operations. Aseptic loosening, including osteolysis, was still reported to be by far the most common cause of revision across all age groups, accounting for over 80% of revisions.
347. Although the average age of patients in Sweden who received total hip arthroplasties after 2001 fell below 70, and for male patients at least, it is similar to the average age of patients in England and Wales who received a Pinnacle Ultamet prosthesis, one cannot justifiably seek to draw parallels between those two groups by reference to mean or median ages. The demographics are too different. There would still be far more elderly patients in the Swedish registry than in the Pinnacle Ultamet cohort in the UK. As with the earlier SHAR reports, and for the same reasons, the CRR for all implants, and the CRR for cemented implants are insufficiently representative to use as reliable comparators even before the inclusion of HXLPE and newer products is taken into account.

348. As Mr Whitwell pointed out, the 2014 SHAR report does show a greater incidence (or predicted incidence) in survivorship of implants over time, irrespective of the means of fixation. That improvement is most evident in the 10 years commencing in 2005, when the new generation of implants had already been introduced, although cemented metal on conventional polyethylene remained the most popular articulation at the start of that decade. The report contains data relating to specific brands of implant. Unlike its predecessors, this includes data for the Pinnacle cup with the Corail Stem, but they are not split by bearing surface, and therefore, as Ms Smith acknowledged, cannot be treated as representative of the performance of the metal on conventional polyethylene liner within the Pinnacle system.
349. The 2014 SHAR report does not call into question the reliability of the findings in the earlier SHAR reports. In terms of the historic figures for 10-year survivorship of earlier cohorts (all ages, all diagnoses and all causes of revision) the 2014 report is largely consistent with the rates reported in the 2000 and 2002 reports. By then the registry had over 20 years' worth of data since 1992 when the surgical techniques for uncemented prosthesis changed, against which the CRR that had been calculated in 2000 and 2002 could be compared. Survival rates for revisions for all causes and all diagnoses from the latest 10-year periods from 1994-2004 and from 2005-2014 were broadly consistent with each other at around 94%-95%, (so the overall CRR had dropped from 7.2% in the 2002 SHAR report, to approximately 5%-6%) with confidence intervals that either overlapped or were within 1% or 2% of each other.
350. However, and despite the report showing an improvement on the figures for 1988-1998 and 1992-2002, the survivorship figures for *uncemented* implants during the period from 1994-2004 remained lower than the figures for all implants over the same period. It was now 91.7% (90.8%-92.5%) – or, in terms of CRR, a point estimate of 8.3% with a range of 7.5%-9.2%. The graph on page 92 of the report illustrates that this group showed a more marked decline than the 2005-2014 group from after around 4 years following the primary operation.
351. It can be inferred that the difference between the 10-year CRR for patients who received uncemented implants in 1992-2002 (in the 2002 SHAR report) and 1995-2004 (in the 2014 SHAR report) is likely to have had some connection with whatever had caused the much higher revision rates in the 1992-2002 data for patients without osteoarthritis whose implants were revised for causes other than aseptic loosening, which presumably had been addressed. However, that cannot be the sole explanation, as the comparative CRR for patients with osteoarthritis and aseptic loosening also decreased.
352. As might be expected, the 2014 report indicated that survival rates for all prostheses implanted in patients with osteoarthritis revised for aseptic loosening, including osteolysis, would be better than those for patients overall. The 1995-2004 cohort had a survivorship of 96.7% (96.6%-96.9%), a CRR of 3.3%, and the 2005-2014 cohort 97.9% (97.6%-98.1%), a CRR of 2.1% across all implants irrespective of fixation and age of the patient. The confidence intervals around those two 10-year sets of figures are close, though they do not overlap. Those figures were also broadly in line with the comparable CRR figures for cemented implants alone, irrespective of the material in the articulating surfaces.

353. The figures for *uncemented* implants in patients with osteoarthritis revised for aseptic loosening show a greater divergence in the survival curves over 10 years between the earlier and later cohorts. The figure for the 1994 group is 95.2% (94.4%-96.0%), a CRR of 4.8%, whilst for the 2005 group it is as high as 98.8% (98.4%-99.1%), a CRR of only 1.2% - 4 times lower than the CRR for the 1994 group, and even lower than the rates for cemented implants. Of course, the data concerning implants in the 10 years from 2005-2014 would include a considerable amount of data relating to metal on HXLPE articulations, even though for the longest periods of data collection within the 10 years observed, contributing to the Kaplan Meier analysis, most of the implants would still have been metal on conventional polyethylene.
354. Thus, according to the data in the 2014 SHAR report, the CRR for the uncemented hips in the 1995-2004 cohort, which insofar as they were MoP articulations would mostly have used conventional polyethylene, were around 8% overall, but only around 5% in patients with osteoarthritis where the cause of the revision was aseptic loosening. Whilst those figures appear to have improved to less than 3% and 1.2% respectively in the 2005-2014 group, there is no overlap in the confidence intervals between the two groups of data, but a clear gap, indicative of a reason for the difference other than random variation, though the gap is not large. There is no obvious explanation for the difference in the report. One possible factor (though by no means the only one) is the increased use of HXLPE in the period from 2005-2014, especially from 2010 onwards.
355. The data over 10 years in the 2014 SHAR report are not broken down into different age groups either. Although, as one might expect, there was some degree of correlation between the statistics for survivorship in younger patients and patients who had uncemented prostheses in the earlier SHAR reports, there is no breakdown by patient age that enables a meaningful comparison to be made with either set of age-related tables in the 2000 and 2002 reports. There are therefore insufficient data to enable an assessment of what, if any, impact the age of the patients would have had on the CRR for the two 10-year cohorts considered in the 2014 SHAR report. Age did make a significant impact on the 10-year survival rates recorded in the 2000 and 2002 reports, including the rates for those patients with uncemented prostheses who had their operations in and after 1992.
356. The age graphs for revisions in the 2014 report, which are confined to all implants and all diagnoses, cover the full 23-year period from 1992 to 2014, and cannot be directly compared with the graphs in the 2000 and 2002 reports. What they do illustrate in broad terms is that over that period, hip implants in patients aged under 50 still have a worse survivorship record than those in patients aged 50-59, and both those groups have a worse survivorship record than implants in patients aged 60-75 and 75 and older. The worst performing group, women aged under 50, have an implant survival rate of 80% at around 12.5 years, whereas in men the rate is 80% at 14 years. For the group aged between 50 and 59, the 80% survival rate is to 15 and 16 years respectively. Therefore, for all the under-60s, at least 20% of hip implants were still failing before 16 years, mainly for aseptic loosening/osteolysis. It was obviously too early to tell what the impact of the introduction of HXLPE would have on the longer-term survivorship data.
357. Unlike the data in the SHAR 2014 report for the period 2005-2014, none of the data from 1992-2002 and very little data from 1994-2004 would relate to newer products.

The impact of HXLPE and newer products on the data for the 2005-2014 cohort, where the CRR for uncemented implants were much more in line with those for patients with cemented implants than they were in respect of any earlier cohort, was likely to have been considerable. For those reasons, I conclude that the figures in the SHAR 2014 report for the decade from 2005-2014 are insufficiently representative of the likely performance of an uncemented metal on conventional polyethylene prosthesis available in 2002 for me to be able to rely on them, and that they should therefore be left out of further consideration.

358. It cannot be deduced from the SHAR 2014 report that 5% is a reliable rate to choose as representing the likely CRR for all metal on conventional polyethylene prostheses, whether cemented or uncemented, that would have been given to the Ultamet cohort in 2002 if they had not received a Pinnacle Ultamet prosthesis. It is obvious that the CRR for a comparable uncemented MoP prosthesis would not have been as low as the 4.8% for the 1995-2004 cohort with osteoarthritis who were revised for aseptic loosening alone. In any event, one must compare like with like as far as is possible, and the NJR statistics relied upon for the CRR of the Pinnacle Ultamet prosthesis relate to revisions of all patients for all causes. The CRR for all uncemented implants in the SHAR 2014 report for the period 1995-2004 was 8.3%, though once the confidence intervals are taken into account, they could have been as low as 7.5%.
359. In summary, in terms of ascertaining what the CRR at 10 years might have been if the patients who received an Ultamet implant had been implanted with a MoP articulation instead, using only conventional polyethylene, the three Swedish registry reports indicate that they would probably have been somewhere between around 7.5% and 15% (the latter figure being a conservative interpretation of the data in the 2000 and 2002 reports). One cannot say precisely where in that spectrum the CRR would lie.

#### **COMPARISON OF THE SWEDISH DATA WITH THE NJR DATA ON ULTAMET**

360. Ms Smith accepted that if one took at face value, and used the point estimate of 13.98% for the 10-year CRR for Pinnacle Ultamet in the most recent NJR report (which relates to *all* causes of revision and includes head/liner exchanges) as a measure of short-term survivorship, the Pinnacle Ultamet performed better than the 10-year revision rate reported in SHAR 2000 for all uncemented implants for the indication of osteoarthritis and aseptic loosening alone, in patients of any age in the 1988-1998 cohort.
361. The SHAR 2000 figure underestimated the CRR for all causes and all underlying conditions by 1%-2%. It also excluded head/liner revisions. To a small, but unknown and incalculable extent those missing factors would be counterbalanced by the fact that there were some older patients within the Ultamet cohort who might otherwise have received a MoP cemented hip, whose predicted revision rate over 10 years would be considerably lower. However, even if this cancelled out the notional increase of 1%-2% completely, the Ultamet's performance would still be better.
362. So far as the younger patients making up the majority of the Ultamet cohort were concerned, Mr Whitwell agreed that if one took the 10-year CRR for the Pinnacle Ultamet from the NJR and compared them with the reported rates for young patients in SHAR 2000 of 20% at 10 years (which are subject to the same limitations), they present a "*significant improvement in terms of survivorship*".

363. The comparison of the data relating to the cohort of patients who were the subject of the tables in the SHAR 2000 report with the data relating to the 1992 cohort of similar patients in the SHAR 2002 report indicates that the surgical techniques used prior to 1992 had no significant impact on the reliability of that earlier data, which appear to me to be robust and reliable.
364. A CRR of 13.98% was also self-evidently better than the 15% at 10 years CRR that Ms Smith and Professor MacGregor accepted as the across the board figure extrapolated from the 2002 SHAR report, which unlike its predecessor did include head/liner exchanges in the definition of “revision,” and was not confined to revision for aseptic loosening. To that extent the SHAR 2002 data is more directly comparable with the data in the NJR. Although a CRR of 13.98% is almost double the 7.2% CRR in the SHAR 2002 report for the most recent period of 10 years recorded in that report for all diagnoses, all revision types, and all means of implant, (i.e. 1992-2002) that is hardly surprising, given that over 90% of the patients contributing to that data were elderly and received a cemented prosthesis.
365. For those patients with *uncemented* implants who were revised for osteoarthritis and aseptic loosening during that period, the CRR in the 2002 SHAR report are quite similar (in the region of 13%) but that is not a like-for like comparison, because the NJR figure for the Ultamet relates to revisions for all causes, and if a notional 1%-2% upwards adjustment is made to the SHAR rates for uncemented implants to take account of other conditions and causes of revision, consistently with Mr Whitwell’s evidence, the Ultamet rates would be better. The figure of 13.98% was far better than the recorded CRR for all patients with uncemented implants in SHAR 2002, which was more than 20%, but those figures are much less reliable, for reasons I have already explained. However, when one takes out the data that appears to be distorting the SHAR statistics, 13.98% is not materially different from the CRR for patients with osteolysis revised for aseptic loosening in that report if one adjusted them by a notional 1%-2% to more accurately reflect the likely CRR for all uncemented implants.
366. Tellingly, the figure of 13.98% is also comparable with or better than the figures for all groups of patients aged 59 or younger in the SHAR 2002 report, irrespective of whether the implant was cemented or uncemented.
367. The comparative exercise based on the SHAR 2000 and 2002 reports suggests that, to the extent that at the end of the 1990s designers, manufacturers and surgeons were trying to find a prosthesis that performed better than the existing MoP prostheses fitted in younger and more active patients, DePuy were at the very least going in the right direction with the Pinnacle Ultamet prosthesis, even though the comparison is only being made over the first 10 years of the life of the implant, and the real aim of introducing the alternative articulations was to tackle the “cliff-edge” phenomenon of failure after that period.
368. The comparison with the data in SHAR 2000 and 2002 indicates that the Ultamet performed at least as well as, and probably better than, the comparator prosthesis would have done or was expected to have done over the first 10 years. That is the only data that would have been available for comparative purposes and informed actual expectations of short-term survivorship at the time when the Pinnacle Ultamet came

onto the market (or in the case of SHAR 2002, within around 6 months after its introduction).

369. So far as the SHAR 2014 report is concerned, a CRR of 13.98% is higher than the CRR for uncemented prostheses in respect of both the later 10-year cohorts from 1994-2004 and 2005-2014. However, the introduction and use of cross-linked polyethylene and HXLPE and newer types of products confounds that data to some unknown extent. It makes the data for the latest 10-year period insufficiently representative to be reliable for any comparative purposes.
370. I am not persuaded that the 2014 SHAR report provides better or more reliable data in respect of the actual or likely performance of a comparator MoP prosthesis in the 10 years after primary surgery than the 2000 and 2002 SHAR reports. However, if one wished to eliminate any possibility that the data for the periods from 1988-1998 or 1992-2002 was significantly confounded by the surgical techniques in Sweden to be insufficiently representative of the CRR for a comparator metal on conventional polyethylene prosthesis, it would be possible to use the figures from the 2014 SHAR report which relate to all uncemented prostheses implanted in the period from 1994-2004, irrespective of the patient's age. That period includes 2002. The underlying data should not be confounded to a significant extent by the introduction of HXLPE or newer products.
371. A straight comparison between the point estimates of 8.3% and 13.98% achieves a difference of only around 5.7%. However, the confidence intervals, which Ms Smith stressed as being far more important than the point estimate, suggest that the CRR for the uncemented MOP prosthesis could have been as high as 9.1%, and the lower level of the confidence intervals around the 13.98% figure is 13.18% even before any adjustments are made to the NJR data for confounders. That means that if an allowance is made for the margin of error around the point estimate, the gap between the two CRR could be as low as 4%. Even if one were to make some notional adjustment for the small number of cemented prostheses that would have been used in some of the older members of the Ultamet cohort, which would have had a lower incidence of failure, one cannot possibly conclude from this data that the Pinnacle Ultamet prosthesis had a materially greater risk of failure than the comparator.
372. Therefore, even taking the figures for the performance of the Pinnacle Ultamet prosthesis from the NJR data in the 2017 report at face value, and treating them as completely reliable, without making any adjustment for confounding factors, I cannot be satisfied on the basis of the Swedish data that there was a material difference between the CRR of the Pinnacle Ultamet prosthesis at 10 years, and the CRR of a metal on conventional polyethylene prosthesis given to the patients in the Ultamet cohort over the same period, if one were comparing like with like.
373. If the Swedish data give a sufficiently reliable indication of how the comparator metal on conventional polyethylene prosthesis would have performed, albeit that it is impossible to carry out a precise comparison, the claimants' claim must fail because they have not proved the abnormally increased risk of early failure on which it depends.
374. If the Swedish data are insufficiently reliable, they are still the best available data from which to gauge the performance of a metal on conventional polyethylene



prosthesis implanted in 2002. For the reasons set out below, the CRR in the NJR based on data on all MoP articulations are far less reliable as a source of information about the likely performance of a metal on conventional polyethylene articulation over 10 years, whether the evaluation of the CRR hypothetically took place at the time when the Ultamet was introduced to the UK market or at any time thereafter. On that hypothesis, the claimants' claim also fails because there is no reliable data on the performance of an appropriate comparator with which the performance of the Pinnacle Ultamet prosthesis can be compared.

#### **4.4 THE NATIONAL JOINT REGISTRY DATA**

375. Ms Smith considered registries which collected data on the clinical use of total hip replacements, and specifically those that collected data in respect of the Pinnacle implants. She was asked to identify those registries which address revision as an outcome measure and which compare Pinnacle Ultamet within other bearing surfaces within the same system and with other bearing surfaces on the market at the same time. She concluded that there were only three registries, namely, the NJR, the Australian Registry and the New Zealand Joint Registry that were suitable to examine the performance of the Pinnacle Ultamet. She used the NJR as her primary source and looked at the data from the other two registries to see if it was broadly consistent, which she concluded it was. I am not persuaded by that conclusion, for reasons I shall explain.
376. The NJR is the largest registry of its kind in the world. Its 14<sup>th</sup> Annual Report, published in September 2017, contains surgical data to 31 December 2016. It notes that the total number of records in the registry are approximately 2.35 million. Ms Smith's evidence was given by reference to the previous Annual Report, but the claimants helpfully set out the figures from the 2017 Report in their written closing submissions, and the 2017 Report was in evidence.
377. Data collection started in April 2003. The data are returned on forms designed by the NJR, which have been revised on various occasions between 2003 and 2007. For the period in which the Pinnacle Ultamet was implanted in patients in the UK, the NJR covered only England and Wales; it has since included data from Northern Ireland and the Isle of Man. The data is completed by the surgeon or a member of the clinical team immediately after the operation, and therefore the form would not be altered in the light of subsequent histological reports.
378. The original form set out a number of different reasons for revision, of which the surgeon could choose one or more, but there was no separate box to tick for ARMD or ARPD. There was a box for aseptic loosening, which surgeons would tick if the revision was for osteolysis, and a box marked "other". A surgeon revising for ARMD could have ticked either or both these boxes. It was only on 1 December 2007 that a new version of the form was introduced with additional details to be filled in. These included a box indicating whether the patient had consented to the data being recorded. The new form also added information on the patient's body mass index (BMI) and changed the options for recording reasons for revision. There was now a box marked "Adverse Soft Tissue Reaction to Particulate Debris". Professor Pandit's evidence, which I accept, was that in practice surgeons interpreted that box as referring to ARMD because soft tissue damage was generally associated with MoM articulations. No version of the form contained information about patient activity

levels, or about co-morbidities beyond the ASA grade (which would indicate the existence of co-morbidities, but not what they were).

379. Ms Smith and Professor MacGregor said they had confidence in the reported outcomes in the NJR because of the high number of participants included in the NJR data, and the fact that registries report on the entire population receiving the treatment, which reduces sampling bias. In 2017, there were 890,681 primary hip procedures recorded of which only 24,103 were reported as having been revised – *under 3% of all prostheses recorded as having been implanted since 2003, irrespective of the articulating surfaces or means of fixation*. That puts the statistics on which reliance is placed into perspective.
380. Ms Smith very fairly agreed in cross-examination that just because a registry is very large in terms of study subjects it does not mean that the data set cannot be subject to bias and confounding, nor does it mean that it will necessarily contain sufficient data to allow an adjustment to be made to control for or eliminate identified confounding factors.
381. In order to use data to identify survivorship of individual prostheses on which to base a survivorship analysis, there are two key factors: compliance and linkability. Compliance (coverage) refers to the proportion of operations in England and Wales that are reported to the NJR. Linkability compares the number of records submitted with the patient's NHS number with the number of procedures recorded in the NJR. The NHS number is required to link all primary and revision procedures relating to a single patient. In the absence of the NHS number, it may be possible to use other information, such as the patient's date of birth, name and address to link the records, but it is far less easy and such information may not be provided.
382. So far as compliance/coverage is concerned, there is no centralised database recording all surgical procedures. The NJR used two other sources of data to calculate the total number of relevant operations undertaken, in order to work out the percentage of those operations that were reported to it. They did this by reference to all hip and knee procedures, not by reference to hip procedures alone. The two sources of data were (1) the number of levies imposed on the sale of prostheses; and (2) the number of operations recorded by NHS databases ("HES" in England and "PEDW" in Wales).
383. Both sources have limitations: a prosthesis which is purchased may not be used in an operation that same year, or at all. That explains why compliance levels measured by reference to levies were sometimes stated as 100% or more. The NJR stopped using levies as a source after 2014. There is also a degree of missing data relating to Pinnacle Ultamet specifically. It was common ground that not all such prostheses sold were recorded in the NJR. Relying on sales data produced by one of DePuy's witnesses, Ms Erin McKibben, Ms Smith identified that only 78.4% of sales of Pinnacle Ultamet were recorded in the NJR.
384. As for the databases, they do not include private treatments, and the coding of HES records was sometimes incorrect. This coding error led to operations which were not revisions being recorded as revisions. Ms Smith agreed that neither of these data sources was ideal. She said there were lots of health warnings in the NJR annual reports about compliance data even as late as 2010. She agreed that one could multiply the compliance rate by the linkability rate to work out the percentage of

usable data but said that was a very rough and ready approach. Her concern about it was that the compliance rate is inaccurate.

385. If records are incapable of being linked, they are excluded from the survivorship analysis. Linkability depended on there being sufficient data available to link the primary procedure with a revision, which partly depended on enough details being supplied at the time of each operation. If a patient did not consent to their personal details being kept, or their consent was not recorded, that could result in the revision not being linked to the primary operation. Ms Smith said that linkability results improved after the NJR obtained a specific exemption under the Data Protection Act; prior to that, there was a lot of missing information on consent.
386. It was common ground that the proportion of usable data has improved over the years, but the level of usable data in the early years of the NJR, especially in the first year of data gathering, 2003-2004, was low. Plainly there was no data available at all for 2002, when the Ultamet was introduced. Ms Smith's initial evidence was that data completeness and linkability of patient records in the NJR was 90% by 2007, but in cross-examination she explained that she was referring to linkability only; that was a reference to a figure given by the NJR itself for the percentage of operations for which a "person-level identifier" was available in 2007. Mr Spencer put to her that if one multiplied compliance by linkability, it was only around 63% in 2006-7, and one only got to 90% in 2010. Ms Smith said that was not the approach she would take, but she agreed that as a rough calculation of how many of the total operations in a year were in the registry and linkable, the figure of 87.29% was produced for the year 2007-2008. On her approach, the percentage of usable data only reached 89.2% in the calendar year beginning January 2007. The two percentages are not that far apart, and both are still below the 90% of usable data that the experts regarded as optimal. Coverage in SHAR, by contrast, was over 90% at all material times, 98% in the 2014 report.
387. For the purposes of the 2017 NJR report, which gathered data to 31 December 2016, the only actual ten-year data that was available was on primary hip replacements performed before 31 December 2006 (and that would only give the recorded revision rates over 10 years for a very limited number of patients who had been implanted during that period). The critical data for the CRR calculations in the 2017 NJR report therefore relates to primary operations carried out during a period when there were significant levels of missing data. All the data contributed to the Kaplan Meier analysis, even if patients had only been followed up for shorter periods than 10 years.
388. On Mr Spencer's rough calculation, the percentage of usable data for that 10-year period was only somewhere between 57% and 63%. This is reflected in the fact that there were only 540 hips at risk for 10 years for the purpose of the 10-year CRR relied upon by the claimants, as Ms Smith accepted. Even the data drawn on for some of the cohorts with less than 10 years' data would have been drawn on usable data of less than 90%. Therefore, by the time of the trial, the NJR data had not yet reached the stage where there was reliable ten-year follow-up from cohorts with at least 90% of useable data. The data used to produce CRR over shorter periods, especially in the period up to April 2010, when the first set of MHRA guidance on MoM hip prostheses came out, would be even less reliable.

389. The fact that data were missing did not necessarily mean that the data source was unreliable; it would be less reliable, but if the missing data were missing largely at random, that would be of less importance than if there were some systematic nature to the missing data, e.g. if they were missing for particular procedures or implants, or particular groups of patients. Ms Smith said that she would be surprised if there were a large amount of bias and the situation was consistent with a general lack of data from different surgical centres rather than from different types of patient. She referred to attempts made over the years by the NJR to ascertain if there were any systemic differences between the characteristics of patients in NHS data that were in the NJR and the characteristics of patients that are not in the NJR, which did not reveal such differences, and to studies that she herself had carried out to like effect.
390. Mr Antelme accepted that the evidence appeared to indicate that patient factors did not seem to be missing on a systematic basis. However, he contended that the missing data still gave rise to some uncertainty because of surgeon factors. The centres that did not return the data may have had surgeons who performed far more (or far fewer) revisions than might be expected, and whose revision rates might have had a distorting impact on the overall data – known as statistical “outliers”. I am bound to say that I am sceptical about this, as the overall acceleration in revision rates, including by the outliers, largely took place in and after 2010, after a time when the data set was much more complete.
391. Both the claimants’ experts stated in the joint expert report that revision rates past 10 years should be used more cautiously, and Ms Smith agreed that it was difficult to be confident that the data was truly missing at random when there was a large quantity missing. Professor MacGregor also agreed that the more data there is missing, the more difficult it is to assess whether what is missing creates a bias or not. As one might expect, the missing data had an impact on the confidence intervals, which narrow over time.
392. The NJR itself expressed some concerns about the impact of missing data. In 2014 it produced a data quality strategy, reflecting on the limitations on the data available to it. It commented: “*without all eligible data though, especially revision cases, the NJR may not be able to provide a clear picture of performance and this is especially important where poor performance is concerned.*” The NJR also said in its 2016 report that “*patients with longer follow-up might be less representative of the whole cohort of patients undergoing primary joint replacement than those patients with shorter follow-up.*” I must bear those caveats in mind when assessing how much weight I am able to place on the NJR statistics relied upon by the claimants.

#### **THE CRR OF THE COMPARATOR PROSTHESES**

393. Like the data in the SHAR 2014 report, the 10-year NJR data to the end of December 2016 includes data relating to various products and materials which came onto the market at the same time as, or after the Pinnacle Ultamet, including products that were introduced after it was withdrawn from the market. The CRR from the NJR data for MoP articulations (including HXLPE) are based on data that have become more complete over time, but the whole period includes data relating to cross-linked and HXLPE articulations, and the NJR data also include information pertaining to products that were introduced at around the same time as or later than the Pinnacle

Ultamet (including the different MoP articulations within the Pinnacle prosthesis itself).

394. For those reasons alone, it is difficult to have any confidence that the CRR recorded in the NJR are sufficiently reliable reflections of the likely 10-year revision rates of a suitable comparator, i.e. a metal on conventional polyethylene implant that would have been given to patients within the Ultamet cohort. The NJR reports contain nothing similar to the graphs in the 2014 SHAR report to give even a rough indication of what percentage of MoP articulations, and articulations generally, had liners made of HXLPE at different points within the 10-year period under consideration.
395. Moreover, the pattern of use of each type of articulation changed over time. I have already referred to the fact that the average age of patients implanted with the Pinnacle Ultamet increased over time, but Ms Smith's analysis did not reflect this. The follow-up periods for different articulations also differed; in the light of the MHRA guidance, from at least April 2010 onwards, MoM articulations were followed up for longer than other prostheses and, at least in some patients, more frequently. It is accepted that this asymmetric surveillance is a potentially confounding factor.
396. On the latest figures from the 2017 Annual Report of the NJR the CRR for uncemented MoP prostheses over 10 years are stated to be 4.18% (confidence intervals 3.97%-4.39%) and for cemented MoP over the same period 3% (2.89%-3.09%). The figures for hybrid MoP implants are somewhere in between - 3.40% (3.20% -3.62%). I have already explained why I consider the figures for uncemented prostheses are the most appropriate to use for comparative purposes.
397. The fact that the NJR data relate to a period after the introduction of the Ultamet liner would not, in and of itself, eliminate the NJR as a reliable source of information about the actual or predicted performance of a comparator MoP prosthesis over 10 years, provided that the comparator was already on the market in 2002. If the NJR had gathered data on metal on conventional polyethylene prostheses and split it by cemented, uncemented, hybrid and reverse hybrid fixations, there would have been a viable argument that the Court (or the experts) could extrapolate figures for comparison purposes that were more reliable than those based on the Swedish data. However, I am not persuaded that it is possible to use the data from the NJR to produce a sufficiently reliable figure for the CRR over 10 years of a metal on conventional polyethylene prosthesis that might have been implanted in the same class of patients as received a Pinnacle Ultamet. There is insufficient information to be able to carry out that exercise, and there is far too much guesswork involved in trying to factor in all the variables and take account of the confounding factors.
398. I have already explained why I have rejected the claimants' submission that the MoP statistics in the NJR can be safely relied upon for the comparison exercise because the introduction of HXLPE would have made no statistical difference to the figures in the short-term. I am not satisfied of the premise. I do not regard it as a legitimate exercise to use the 9-year NJR data relating to the relatively small number of metal on conventional polyethylene Pinnacle implants even as a means of checking the reliability of the CRR based on the wider MoP data in the NJR. The Pinnacle system was a new product, designed to produce minimal articular wear, and if orthopaedic surgeons decided to use the conventional polyethylene liner in that system, rather than one of the alternative harder bearing surfaces, it would have been used for their older,

less active patients. One cannot make a reliable comparison between how the wider group of metal on conventional polyethylene prostheses and metal on HXLPE prostheses would perform, based on data pertaining to the Pinnacle system alone.

399. I cannot accept the NJR data as being a more reliable source than the data in SHAR for the purposes of comparing the short-term survivorship of comparator bearings with the Pinnacle Ultamet, notwithstanding that SHAR is a smaller registry. SHAR is a better source of data in respect of the likely performance of metal on conventional polyethylene, since the data in SHAR clearly indicates that at least for the 10-year periods from 1992-2002 and 1994-2004, most if not all the implants would have used articulations with conventional polyethylene rather than HXLPE.
400. In my judgment, insofar as it is possible to make any comparative evaluation at all, the SHAR data is likely to provide a more reliable indication of the 10-year performance of an uncemented metal on polyethylene prosthesis that would have been implanted in a patient falling within the Ultamet cohort in or around the time when the Pinnacle Ultamet was introduced onto the market, even though that data provides a spectrum of possible CRR for such articulations.
401. I am not persuaded that the 10-year CRR for a comparator MoP prosthesis or prostheses implanted in the patients in the Pinnacle Ultamet cohort would have been as low as 5%, let alone the 4.18% submitted by the claimants based on the figure for uncemented MoP articulations in the table in their final written submissions (with a confidence interval of 3.97%-4.39%). The figures in the NJR for MoP are not reliable for this purpose, not just because of the mix of conventional and cross-linked and HXLPE and the inclusion of later products in the statistics, but also because most patients who were actually given MoP hips, certainly at the start of the 10-year period, would have been the older, less active patients, many of whom would never have been candidates for receiving a Pinnacle Ultamet prosthesis. Despite Ms Smith's evidence that age made no appreciable difference to the NJR statistics, she accepted that there was no reliable means of determining what impact activity levels would have had on them.
402. However, in fairness to the claimants I will go on to consider the NJR data pertaining to the performance of the Pinnacle Ultamet on an assumption in their favour, contrary to my findings above, that the NJR figures *are* a reliable indicator of the likely 10-year survivorship of a metal on conventional plastic articulation implanted in a patient who would otherwise have received a Pinnacle Ultamet in or after 2002. I will use for these purposes the point estimate of 4.18% for uncemented MoP prostheses (and the confidence intervals around it) in the NJR report for 2017.

#### **THE CRR OF THE PINNACLE ULTAMET PROSTHESIS**

403. I have already referred to the fact that the Pinnacle Ultamet CRR relied on by the claimants by way of comparison, taken from the 2017 NJR report, are a point estimate of 13.98% with confidence intervals of 13.18% to 14.83%. Although the percentages in the 2016 report (and, indeed, earlier NJR reports) are slightly higher, Ms Smith said, and I accept, that those differences could be ascribed to random variation. It is important to bear in mind that these data include all head sizes, not just the 36mm. I have already noted that they are considerably better than the figures for all MoM articulations.

**EFFECT OF THE FINDINGS OF OTHER REGISTRIES AND STUDIES**

404. The Australian Joint Registry report relied on by Ms Smith and Professor MacGregor by way of support, as a secondary source said to be consistent with the CRR calculated on the basis of the NJR data, does not compare the Pinnacle Ultamet with external comparators; instead, there is an intra-Pinnacle comparison between the different articulating surfaces, broken down by different stems and head sizes. Ms Smith relied on the 2016 Report, but the figures in the 2017 Report are statistically consistent. The intra-Pinnacle comparison excludes head sizes of 32mm or below, and in the case of the S-ROM stem, the data is confined to revisions where the primary revision was for osteoarthritis. The confidence intervals are far wider than those in the NJR tables.
405. For the large head MoM Pinnacle prostheses using a Corail Stem (of which there were fewer implants reported in the Australian registry than the NJR) the point estimate is 12.9% with confidence intervals of 10.5% to 15.9%. For the Summit stem the point estimate is 8.8% with confidence intervals of 6.8% to 11.2%, and for S-ROM, 6.9% with confidence intervals of 4.4% to 10.7%. There is a notable variance between the CRR for the different stems, which is rather odd if the main reason for early revision was ARMD. One would expect the incidence of ARMD to be broadly similar across the range. Not surprisingly in the light of the width of the confidence intervals, the different categories overlap at one or other end of the range. One knows nothing about any confounding factors in Australia.
406. Ms Smith very fairly pointed out that the highest revision rates for the Ultamet recorded in the NJR were with the S-ROM stem, whereas in Australia they are the lowest (albeit that the data is circumscribed in the manner I have mentioned). The claimants accepted that the data relating to Pinnacle Ultamet prostheses using S-ROM stems in the NJR report are affected by the “outlier” issue, addressed below as one of the significant confounding factors. There are too few Summit stems recorded in the NJR to yield meaningful data about them. Therefore, only the Corail figures in the Australian registry can be directly compared with the NJR figures. The range in the NJR of 13.18%-14.83% for all stems unsurprisingly fits within the much wider Australian range for the Corail stem of 10.5% -15.9%. Ms Smith seeks to rely on the consistency of the findings between these two datasets for this particular combination, leaving out of account the much lower figures for the other combinations of stem and cup, despite the fact that the latter were more popular in Australia.
407. Mr Antelme drew attention to the fact that Ms Smith’s own table of the CRR for the different articulations within the Pinnacle system (based on the NJR 2016 Report) indicated a CoC point estimate of 4.62% with confidence intervals of 3.92%-5.42%, and yet the rates for the Pinnacle CoC articulation across all the different stems in the Australian registry went up to 8.5%. Whilst that does raise some concerns about consistency, there are all kinds of reasons why the Australian experience with survival of CoC articulations might have been different from that in the UK, and one really cannot speculate about them in the absence of evidence.
408. The fact that Ms Smith and Professor MacGregor were prepared to accept the data in the Australian registry as supportive of the CRR in the NJR troubles me somewhat. There appears to be a degree of selectivity involved. The S-ROM data, for example, suggested that the performance of the Ultamet is much closer to that of other types of

articulation within the Pinnacle system than the NJR data suggests, even if one were to increase the figures by a notional 1%-2% to compensate for the fact that it is confined to patients with a primary diagnosis of osteoarthritis. The much wider confidence intervals in the Australian data also invite caution. Overall, I am not persuaded that the Australian data provides reliable support for the NJR figures, particularly since the Australian data is confined to the intra-Pinnacle comparison which I have already held to be the wrong comparison for the purposes of evaluating the entitled expectation of safety.

409. Both Ms Smith and Professor MacGregor also relied on the New Zealand Joint Registry as an additional secondary data source on the performance of the Pinnacle Ultamet, consistent with the NJR. Again, I do not accept that it is sufficiently reliable to be used for that purpose. It contains no data on MoP articulations, because the Registry does not keep separate data on MoP from CoP articulations. It does not use the same Kaplan Meier methodology, but the patient time incidence rate (PTIR) per 100 component years. The PTIR is a standardised measure of revision that takes account of the different lengths of time that patients have been observed. It represents the *average* rate of occurrence of revision during the period of observation in which the patients could be considered at risk.
410. Professor MacGregor described the PTIR as “*less informative when the rate of occurrence of revisions over time is not constant over time*” and as a “*less sensitive indicator of change in the risk over time than CRR*”. The rate of revisions of the Pinnacle Ultamet was not constant over time, as the NJR data itself demonstrates from 2010 onwards. The follow-up periods for MoM prostheses were also different from other articulations. Ms Smith explained that if the PTIR per 1,000 patient years for ARMD was 9.2 it would mean that if one observed 200 patients each for five years, 9 of those 200 procedures would have been revised for ARMD in that time, but the figures would change if the time observed was different. Given the different periods of observation for the different articulations, which Ms Smith accepted to be a potential confounding factor, I am not prepared to accept this data as reliable independent support for the CRR in the NJR, despite the views of the claimants’ experts.
411. DePuy relied on the Nordic Arthroplasty Registration Association data as reported in a study by Varnum and others in 2015, but again that used PTIR. Ms Smith accepted that that study suggested that over a six-year period there was no statistical difference between the CRR for a Pinnacle Ultamet prosthesis and a MoP prosthesis (as with the NJR, there was no distinction in the data between any different types of polyethylene being used). However, the number of prostheses observed was relatively small, and the period of observation was only from 2002-2010. Most importantly, the follow-up period was short. Again, I do not consider that the information in this study is sufficiently reliable to enable any meaningful conclusions to be drawn from it.
412. Mr Spencer asked Professor MacGregor about some studies that had been carried out on CRR for the Pinnacle Ultamet. Professor MacGregor rightly emphasised the caution that is necessary when looking at studies carried out in one particular centre or with limited sets of surgeons, as those factors can introduce bias. He was firm in his view that the NJR was a more reliable source. However, he accepted that one of the studies, by Atrey in 2016, was a well-designed study. It described itself as the largest study of 36mm head MoM Pinnacle prostheses using a Corail stem, and had the



longest follow-up over 9.7 years, with a median of 7.2 years. That study found survivorship of 92.8% at ten years with confidence intervals of 91.6% to 94%.

413. DePuy did not rely on this or any other study as providing data (or CRR) that were more reliable than the NJR data, but they submitted that the other studies produced results that were within the statistical range on which the claimants themselves relied, if and insofar as the claimants sought to rely on the findings of the Australian and New Zealand registries as reliable secondary sources. A point estimate of 7.2% is certainly within the range reported in the Australian registry for the S-ROM stem, but it is outside the range in that registry for large head MoM prostheses with a Corail stem. However, I am not persuaded that this study or any of the others referred to by DePuy, including its own PIN study, which had very high levels of loss to follow-up, take matters any further.

#### **COMPARISON OF THE CRR OF ULTAMET AND COMPARATOR PROSTHESES**

414. On the face of it, the recorded 10-year CRR for the Pinnacle Ultamet Prosthesis in the NJR are more than double the CRR in the NJR for all external comparators, including the MoP articulations which include those with liners made from HXLPE. Indeed, as Mr Oppenheim pointed out, the point estimate of 13.98% is over 3 times greater than the 4.18% point estimate for uncemented MoP articulations in the 2017 NJR report. That differential realistically represents the high water-mark of the claimants' case. However, even assuming the 4.18% figure to be a reliable indicator of the likely 10-year risk of revision of the external comparator if it had been implanted in the Ultamet cohort, the next question is, how much reliance can be placed on the figure of 13.98% (or the confidence intervals around it) as a measure of the risk of revision of the Ultamet over that period?
415. It was not DePuy's case that instead of the CRR for the Pinnacle Ultamet being 13.98% (or somewhere in the range indicated by the confidence intervals) it was in fact some lower percentage. They cannot provide an alternative figure based on working out how the various confounders impact on the 13.98% figure, or whatever figure the Court uses as a starting point for comparing the Pinnacle Ultamet prosthesis with non-MoM prostheses, because no-one has the material with which to do so. DePuy's case is that the confounding features lead to the inexorable conclusion that the CRR in the NJR are not a reliable indication of the rate of early failure of the 36mm head Pinnacle MoM prostheses.
416. Mr Antelme submitted that the Court should not use 13.98% as the starting point for making the comparison, given that the CRR was somewhere within a range, of which the lowest point based on the NJR data was 13.18%, and in the Australian data for Corail stems used with a large head Ultamet, the lowest point in a substantially wider range is 10.5% (or 4.4% across all the different stems). I have already explained why I am not persuaded that the Australian Registry provides anything other than weak support for the reliability of the CRR recorded in the NJR, though I appreciate that some of the figures in it are more favourable to DePuy. Whilst there is some force in the point that in order to test the claimants' case, one should take the lower figure in the confidence intervals rather than the point estimate, because that lower figure could represent the true figure, I will start by assuming in their favour that the CRR is the point estimate of 13.98%. I will test its reliability in the light of the various confounding factors identified by the parties.

### **CONFOUNDING FACTORS**

417. Various potential confounding factors were identified which go to the reliability of the data in the NJR (for the purposes of the comparative exercise) and the question whether there is a reliable or sufficient basis for making the comparison. The claimants' final written submissions concentrated on looking at each of these factors individually and submitting that taken alone, they could not account for the substantial difference between the CRR calculated for Pinnacle Ultamet prostheses and the CRR for other articulating surfaces. In his oral final submissions, Mr Oppenheim emphasised that the claimants' case was that, taken collectively or in any combination, the identified confounding factors could not account for the difference between 4.18% and 13.98%. The CRR for the other articulating surfaces are broadly comparable with each other, whereas the CRR for the Pinnacle Ultamet is much higher.
418. However, the issue before the Court is not whether an individual factor or cumulative factors could account for the reported differences in the CRR as recorded in the NJR report, but whether the CRR and the comparison relied upon by the claimants are sufficiently reliable for the Court to conclude that the risk of early failure for the Pinnacle Ultamet prosthesis was materially greater than that for comparator prostheses.

### **Age and sex**

419. The first accepted confounding factors were age and sex of the patients. It was accepted that uncemented prostheses and prostheses implanted in younger patients had higher CRR than cemented prostheses. However, these confounders alone would not account for the differences in the CRR recorded in the NJR report. Ms Smith carried out a flexible parametric survival model to estimate revision rates for exactly the same types of patients to reduce the risk of bias for these factors, among others, and concluded that although there is some impact, the revision rates for Pinnacle MoM Ultamet implants with Corail stems (the only group on which there was sufficient data to perform the analysis) were still significantly higher than those for other bearing surfaces. She said in cross-examination that age has some impact, but it does not make a huge difference.

### **Body Mass Index (BMI)**

420. It was common ground that BMI is poorly and incompletely recorded and that it is the one data field in the NJR where there is evidence of systemic bias. That is because there will be an under-recording of BMI when the BMI is in the healthy to overweight range. It is much more likely to be recorded when it is obese or morbidly obese. Ms Smith did not think its absence would confound the comparisons to any important degree, though she said that it was something that would perhaps benefit from more research. Whilst I agree that more research would be desirable, it seems to me to be difficult to make the assumption that BMI would make no appreciable difference, in the light of the engineering evidence about the role that loading, as well as other complex and variable factors which will vary from patient to patient, will have on the wear of a bearing.

## Activity

421. The next confounding feature identified by DePuy was activity, though this was controversial. Although both Ms Smith and Professor MacGregor understandably pointed to the fact that there have been no epidemiological studies directly measuring the impact of certain types of activity on the rate of revision of hip prostheses, I am satisfied from the evidence of the generic orthopaedic experts and the unchallenged evidence of Professor Fisher that different types of activity can and will have a significant impact on the amount of wear generated, and that in general terms heightened activity is likely to increase wear rates and thus is more likely than not to accelerate the rates of revision.
422. SHAR 2000 described the young and active patients as the “supreme challenge”, a view endorsed by the exhortation in SHAR 2002 to continue the search for harder wearing articulating surfaces. Mr Whitwell and Professor Pandit both agreed that activity was a critical factor in the survivorship of a prosthesis. Professor MacGregor confirmed that the suggestions in guidance issued by NICE in 2014 that greater activity impacts on the need for revision came from surgeons, from the experience of an assessment group and from discussion in the orthopaedic community. Whilst the link may not yet have been proved to the satisfaction of an epidemiologist, I am satisfied that in real life there is a link between activity and rates of wear and that wear rates will increase with greater activity.
423. The *extent* to which activity will impact on the rates of revision is not something that has yet been the subject of sufficient scientific study – it is not an easy thing to measure. There are numerous complex factors affecting loading on the hip joint even before individual factors such as weight or BMI are taken into consideration. Professor Fisher pointed out that jogging or running increases the loading and the wear produced on all bearings, but it may only account for 10% or 20% of overall activity. In the 1990s, Dr Haas told his patients not to jog. Likewise, climbing stairs increases the loading, but that is something that may happen less often than walking around the house. Individual patients undertake different types of activity with different levels and intensity; a young active patient may undertake double the number of loading cycles on the implant per year than those undertaken by an elderly inactive patient. The information from SHAR provides some support for the view that revision rates are likely to be higher in younger and more active patients.
424. Ms Smith agreed that the nature and level of activity were *potentially* confounding factors, but no data had been collected by the NJR on these matters. She had no data to be able to analyse the confounding influence of activity. She said that age and health status could be used as partial proxies, but they are not perfect measures because some patients will not fit the age and health profile. It was common ground between the orthopaedic experts that there are no good surrogates for activity in other data that has been gathered. As Dr Haas suggested, and Mr Whitwell confirmed, in broad terms activity does not correlate with age in patients aged between 55 and 70, but it may do so outside that range, though there will still be exceptions. There may be 80-year olds who run marathons and paraglide, and very inactive 50-year olds, particularly if they have co-morbidities, or if whatever has caused the need for revision has restricted their activity levels. The mismatch between age and activity is accentuated when prostheses are selected on the basis of activity rather than age. Mr

Whitwell agreed with Dr Haas' evidence that age and activity are both factors that are relevant in the selection of a prosthesis.

425. Mr Oppenheim submitted that whilst in principle activity was a potential confounder, there was no real evidence that activity was a significant confounding factor so far as the NJR data were concerned, because there were no substantial differences in terms of revision rates between the groups of patients receiving CoC (the hardest-wearing articulation) and MoP hips by age band, and that is what one might expect to see if activity made a significant difference to early revision rates. The NJR tables relating to all prostheses and all patients, indicate that the 10-year CRR for CoC are 4.62% with confidence intervals of 3.92%-5.42% whereas the CRR for MoP are similar or better - 3.92% with confidence intervals of 3.33%-4.60%. However, the CoP data is the lowest of all, 2.80% with a range of 2.17%-3.63%. That is a hard on soft articulation.
426. On the face of it, there are no substantial differences in terms of revision rates between the groups of patients receiving CoC and MoP hips irrespective of age. The data overlap between CoC and MoP suggests that CoC has conferred no appreciable benefit over the less hard-wearing polyethylene liner, which is not what one would expect to see, even in the short term, subject to the impact of the cross-linked and HXLPE products contributing to the data on MoP articulations.
427. However, Mr Oppenheim's submission depends upon an assumption that one is comparing like with like, and that may not be a fair assumption. The precise make-up of the different patient cohorts is unclear. If the MoP articulations were mainly offered to the least active patients within each age band, and the hardest articulations, including CoC, were mainly offered to the most active, there is nothing particularly surprising about the rates of wear being comparable. One would expect the harder articulations to be more resistant to greater wear, and the softer articulations to be subjected to less wear.
428. As I have already noted, the comparison between CoC and MoP is further complicated by the fact that the latter will include HXLPE, which could well mask differences in wear and revision rates due to activity. It is reasonable to assume that the alternative bearings, including HXLPE and CoC, might be offered to more active patients, as that was the very reason for their development. The NJR data stratified by age suggests that CoC was mainly offered to the youngest (and thus most active) cohort. However, as DePuy pointed out, the ability to provide a larger head, at least in the earlier years, might also have prompted a surgeon to choose a MoM articulation for an active patient instead of one of the alternative hard bearings, which means that the numbers of active patients falling within the datasets for the latter will not represent the full cohort of active patients. There were concerns about fractures in CoC, at least until the fourth generation was introduced, and that was long after the Ultamet came onto the market. This might have inhibited surgeons from implanting a third-generation CoC articulation in a more active patient. Mr Whitwell was very chary about using CoC.
429. The closeness of the figures for both types of ceramic articulation and the figures for MoP could be explained on the basis that either the more challenging patients have not been offered MoP articulations (which would improve the revision rates for MoP) or that the MoP and CoP articulations include a substantial number of HXLPE liners,

or both. I have already explained why I am not persuaded by Ms Smith's intra-Pinnacle analysis of the performance of conventional polyethylene liners and HXLPE in the short-term that one can reliably assume there is no significant difference in terms of survivorship rates of those materials over ten years.

430. I therefore conclude that activity cannot be dismissed as a potential and significant confounding factor; however, the extent to which it is likely to confound the data is unknown and unpredictable. One cannot draw any reliable conclusions from the fact that age alone was found by Ms Smith to have little impact, or from the fact that the CRR recorded in the NJR relating to CoC articulations are similar to the CRR relating to MoP articulations, irrespective of patient age.

### **Asymmetric Surveillance**

431. Mr Whitwell's evidence was that in his own private practice he recommended follow-ups with X-rays at one, five and ten-year intervals in order to try and detect osteolysis (since he used MoP implants in his primary operations). NHS patients would not be monitored beyond a standard 1-year review, so they would not be seen again unless they were symptomatic and referred by a GP. He said he would have liked patients to have more regular follow-ups with X-rays which would pick up signs of osteolysis earlier. He accepted that if there had been a similar system of review for osteolysis for MoP hips as there was for MoM from 2010 onwards, there would have been some earlier revisions of MoP hips. If that had happened, the CRR figure of 4.18% for MoP hips would have increased.
432. The effect of the enhanced and more regular surveillance of MoM hips of course meant that it was possible to pick up genuine cases of ARMD earlier than would otherwise have been the case, and to effect precautionary revisions before the damage became too extensive. That is not a matter of which complaint can be made. However, the dissimilar monitoring of MoP hips meant that cases of osteolysis, or other reasons for revision which might have been detected much earlier in comparator prostheses were not detected, which would confound the statistics to an unknown extent. Ms Smith accepted that there was potential bias and confounding in relation to asymmetric monitoring when considering reasons for revision as a whole. Again, the impact of this confounding factor cannot be calculated because there is no data that would enable the calculation to be carried out.

### **The increase in revision rates in 2010**

433. Ms Smith accepted that in 2010 there was an increase in the CRR for the Pinnacle MoM hip implants which was not seen for other bearing surfaces. The increase related to a calendar date rather than time elapsed since implantation, and therefore was atypical. The increased rates of revision became apparent soon after the introduction of the MHRA guidance, and with it, the increase in monitoring of patients with MoM implants. Ms Smith did not disagree with the graphs in Professor Scharfstein's reports setting out the position; her own graphs in her second expert report took a slightly different point in time, but she confirmed in her oral evidence that they demonstrate broadly the same change as Professor Scharfstein described. Although the graphs do not show a single large spike in the rates of revision at a specific point in time, they do indicate a significant increase in the number of revisions in the period after the MHRA guidance came out. That was not a coincidence.

434. Ms Smith produced a table (based on the figures from the 2016 NJR report) showing that the 5-year CRR based on data for revisions prior to April 2010 was 2.85% with confidence intervals of 2.34% to 3.44% but in the 5 years from April 2010 to August 2015 the 5-year CRR was 5.25% with confidence intervals of 4.87%-5.66%. She pointed out that this analysis was complicated by the fact that there were relatively few MoM implants followed up for a reasonable period prior to the MHRA guidance. Only 458 Pinnacle MoM hips had been followed up for at least 5 years, and only 97 had been followed up for at least 6 years.
435. I am satisfied that there was a significant increase in the revision rates in and after April 2010, and that the increase was directly attributable to the impact of the issue of the MHRA guidance on surgeons and patients alike. That impact included the panic about MoM hips that was engendered in consequence of the increasingly hysterical media reporting (I was referred to and read a substantial number of the reports, which I am satisfied were a fair representative sample). Since some patients would have presented with symptoms leading to revision during this period in any event, it is impossible to evaluate how much of an impact these factors had on the revision rates, but it was obviously significant.
436. DePuy quite properly accepted that, to the extent that the MHRA guidance reflected clinical requirements to treat MoM hips differently from other articulations, this factor in and of itself does not undermine the comparison between the different prostheses. What *does* undermine the comparison is the impact of the adverse media reporting and the ASR recall on the decisions to revise. I have already referred to the unchallenged expert evidence about the “nocebo” effect. Whilst Mr Whitwell said he was able to provide assurance to his alarmed patients, other surgeons may not have been as successful as he was in doing so. In any event, reported symptoms of pain (including somatic pain) would have influenced both Mr Whitwell and other surgeons in considering whether to recommend revision. They would have taken at face value what their patients were telling them about pain, unless the patient’s description of where the pain was coming from was obviously at odds with there being a hip-related problem. The fact that pain is somatic does not make it any less real an experience for the patient.
437. The MHRA guidance set out a minimum threshold for surveillance which was designed to try and identify the patients who had a problem with ARMD as soon as possible. Once the investigations were complete, the surgeon would give consideration as to whether to revise, or to continue to monitor the patient. Some centres around the country and some surgeons thought that the MHRA set the threshold too high, and therefore ordered cross-sectional imaging or other tests for patients whose blood ion levels were in excess of 4ppb or even lower. Some centres investigated all patients with MoM hip prostheses regardless of whether they were symptomatic or asymptomatic.
438. The further confounding factor that the increased level of monitoring led to, was the adoption by many surgeons of a lower threshold for revision than the MHRA had envisaged and in some cases, a mis-diagnosis of ARMD or a revision that was stated to be for ARMD when in fact it was for unexplained pain or because the patient had a MoM hip, just to make sure that the patient did not go on to develop ARMD. In all such cases, the revision of the hip could not be equated with the clinical failure of the prosthesis. On revising for suspected ARMD in and after 2010, the surgeon would

tick the box marked “ARPD” on the NJR form, but the form would be sent to the NJR before the results of any histopathology tests on the excised tissue were returned, and therefore if the tests indicated that there was no soft tissue damage, and no histological signs corroborative of the clinical diagnosis, there would be no opportunity for correction of the data.

439. Of course, orthopaedic consultants had a real problem in diagnosing ARMD because of the lack of clear diagnostic criteria. Unlike osteolysis, the signs would not show up on conventional X-rays. The expert radiologists confirmed that if a patient with a MoM prosthesis was referred for investigation, the first thing that a radiologist would be looking for would be signs of ARMD. It was more difficult to pick up pseudotumours using ultrasound, though an experienced radiologist who was specifically looking for symptoms of ARMD should have been able to do so. The ability to find a pseudotumour improved with the development of MARS MRI scanning.
440. The consultant would have to place a great deal of reliance on what the patient said about his or her symptoms and about the nature and location of any pain, especially if the test results were equivocal. Mr Whitwell’s view about the reluctance of surgeons to revise on the basis of unexplained pain alone, and the importance of excluding alternative problems such as bursitis or tendonitis, was not borne out by the evidence in the lead claims.
441. There will always be a degree of subjectivity involved in the decision to revise. One surgeon might decide to revise the patient’s MoM hip in circumstances where another surgeon would have decided to wait and monitor. That does not mean that either of them was acting unreasonably, because there always will be situations in which there is a spectrum of reasonable clinical decisions that could be taken. Nevertheless, the decision to revise might have been premature, or turn out to have been unjustified, when all the clinical evidence was examined, as in the case of Mrs Blake. In other cases, if the MHRA guidance were followed, the patient might still have been revised, but the revision would have occurred months or even years later, and that would also have some impact on the Kaplan Meier analysis.
442. It was quite clear on the evidence that, given so many uncertainties about what might happen to the soft tissue in future, surgeons would revise a MoM hip in the light of symptoms that would not lead to a revision of an alternative articulation, because of the general attitude that it was better to be safe than sorry. If the revision was a relatively straightforward exercise, as it would if there were just a head or liner swap, this would be a further factor weighing in the balance in favour of revising.
443. In the wake of concerns about ARMD, it is hardly surprising that surgeons felt it would be prudent to act sooner rather than later; Mr Whitwell’s approach of “if in doubt, revise” (in the context of imaging which revealed a collection around the greater trochanter) was not a maverick stance. It was understandable given his experience from working in a tertiary centre of what could happen to a patient with serious soft tissue damage; some of the surgeons in the lead claims had similar experiences before they operated on the claimants. Two of the surgeons who gave factual evidence in the lead claims, Mr Dunlop and Mr Herlekar, had a very low threshold for revision, a third, Mr Marsh, revised simply because a patient with a MoM prosthesis was experiencing mild discomfort, the fourth, Mr Nargol, was one of

the “outlier” surgeons with an abnormally high rate of revisions discussed in the next section of this judgment. All four of them failed to follow the MHRA guidelines.

444. As DePuy rightly accepted, the fact that surgery may turn out to have been unnecessary does not mean that the decision to revise was unjustified, let alone negligent. Even though there were occasions in which it transpired that the revision surgery was unjustified, it has never been DePuy’s case that the difference in the CRR reported in the NJR for Pinnacle Ultamet and MoP prostheses is entirely attributable to unnecessary revisions for ARMD. DePuy’s argument was simply that where certain factors impact on the decision to revise one type of prosthesis at a given point in time which have no impact on the decision to revise the comparator prosthesis, the comparison will be undermined to an unknown and incalculable extent, adversely affecting its reliability.
445. Here, there were a number of inter-related factors which had an impact on the revision rates; chiefly, the impact of the panic engendered by the adverse media reporting on patients and surgeons alike, and the lowering of the threshold for revision below the MHRA guidelines. Mr Antelme submitted that the resulting increase in the revision rates cannot be regarded as a sound basis for assessment of the short-term survivorship, let alone the safety, of the Pinnacle Ultamet prosthesis. I agree.

#### **The impact of “outlier” surgeons**

446. Ms Smith agreed that a “frailty model” analysis carried out by Professor Scharfstein indicated that there is heterogeneity in the revision rates between surgeons. She said that he demonstrated that there was a statistically significant variation between the surgeons who performed revisions of the Pinnacle Ultamet, but that his analysis did not reveal how much of this was due to random variation. Therefore, Ms Smith carried out a “funnel plot” analysis, which is the technique used by the NJR to identify individual surgeons with higher than expected revision rates. As she explained, a funnel plot is able to distinguish between a variation which is random in nature and that which is statistically significant. The analysis has confidence limits that are similar statistically to 95% confidence intervals, though they are calculated in a different way.
447. Ms Smith’s funnel plot identified four surgeons whose revision rates were statistically relevant because they were significantly above the rates of other surgeons and corresponded with the NJR’s classification of “alarm” calling for investigation. Two of these surgeons, Mr Tulloch and Mr Nargol, were operating in hospitals run by the North Tees and Hartlepool NHS Foundation Trust. Mr Nargol was the surgeon who operated on Mrs Stalker, one of the lead claimants. He implanted a large number of prostheses with S-ROM stems in his patients, though he used a different stem for Mrs Stalker.
448. An “outlier” surgeon is not necessarily a bad surgeon – his higher revision rates may reflect the fact that he had a higher number of patients with complicated medical histories who are more likely to have had worse outcomes. However, if that surgeon is not working in a specialist centre like Mr Whitwell, and he is carrying out twice, or almost twice, the number of revisions that surgeons elsewhere are carrying out in the same period, that does give rise to some concern about the reliability of the statistics. His much higher rates of revision may signify that he has a very low threshold for



revision, or possibly that there was some specific problem with the primary surgery – a difficulty in surgical fixation, for example.

449. As the comparison exercise assumes that revisions within 10 years for reasons other than failure for ARMD or ARPD/osteolysis are broadly similar across the board for all prostheses, irrespective of articulating surface, a problem of that nature could distort the comparison, as it would not be a reason that applied across the board. Either way, the outlier's vastly increased revisions would not be an accurate reflection of the rates of early failure or a sound basis for predicting the likely survivorship of implants in patients overall.
450. Ms Smith's funnel plot also revealed three "outliers" who had *lower* revision rates than other surgeons. She said that statisticians would not remove any of the outliers for the purposes of evaluating the data, but that in order to be consistent, one should remove all outliers if one removed any (an exercise she said she had carried out, though she did not put her analysis in her report and provided the figures from recollection in the witness box). I understand why Ms Smith took the view she did. However, I have been invited to treat the CRR as a reliable basis for assessment of the risk of early failure of the implants for ARMD (on the assumption that the risk of failure for other reasons is broadly the same across all products). In that context, the relevance of the outliers who were performing far more revisions than others, is that it demonstrates the extent to which even a small number of surgeons could produce data which lead to the calculation of artificially *inflated* CRR.
451. There was no evidence that any surgeons were failing to revise MoM prostheses despite there being indications for revision. Indeed, the surgeons with lower revision rates may well have been the ones who were adhering to the MHRA guidelines. The data relating to the surgeons who had lower revision rates illustrate the performance of the prostheses for these purposes, without any suggestion of inappropriate bias. Therefore, I do not accept that the lower "outliers" confound the data for the purposes for which it is being used, in the way that the higher outliers undoubtedly do.
452. The impact of the four outliers who had higher revision rates is telling. They undertook up to 69% of all the Pinnacle Ultamet prostheses used as primary implantations in the earliest years of the NJR. Data from patients implanted in those years was particularly important for the 10-year CRR relied upon by the claimants. Overall, the four surgeons accounted for around 10% of the total number of prostheses implanted, but just under a fifth of the total revision rates identified by Ms Smith. On her figures (which were based on the slightly higher CRR in the 2016 NJR report, as she did not have the 2017 report when she did the calculations) the 10-year CRR for the Pinnacle Ultamet for those surgeons alone were 26.02%, whereas for the non-outlying surgeons the CRR were 12.94%. Ms Smith said, though she did not provide evidence of her workings, that if one counterbalanced this by taking into account the outliers below the confidence limit, the 10-year CRR for the 4 surgeons would go down to around 18% and the CRR for non-outlier surgeons would be around 13.3%.
453. The effect of removing the 4 upper outliers was therefore to reduce the 10-year CRR for Pinnacle Ultamet by as much as 2.6% (on the point estimate alone). In terms of the confidence intervals, Ms Smith's range for a 10-year CRR was reduced from 14.30%-16.87% to 11.81%-14.17%, with a point estimate of 12.94%. I was not supplied with

the figures which would show the impact on the 2017 CRR, which were lower than the 2016 figures, but assuming the percentage impact to be in the same range, one would no longer be looking at a point estimate of 13.98% but something in the order of 11.2%. The figures would be even lower when the confidence intervals were adjusted.

454. Whilst it is possible to calculate the impact of this particular confounding factor on the calculation of the CRR, it is not possible to calculate the impact of all the others I have identified.

### **CONCLUSIONS ON THE NJR STATISTICS**

455. This leads me to the following conclusions:

- i) I cannot safely rely upon 13.98% (or even the lower of the confidence intervals of 13.1%) as being representative of the CRR of the Ultamet prosthesis over 10 years. There are too many potentially confounding factors that could have a statistically significant impact on the CRR. The outlier effect alone would suggest that a CRR of somewhere between 10% and 11% was more accurate, but that is only the start of the adjustments that would need to be made before one could reach any reliable prediction of the CRR within 10 years of implantation.
- ii) The NJR data are incomplete, and there is very little actual long-term evidence relating to the performance of Pinnacle MoM implants. As the NJR itself acknowledged, this in and of itself raises some concerns about the reliability of the data as a measure of performance. Over time, as more data is gathered, the statistics will become more reliable.
- iii) Apart the effect of age and the outlier surgeons on the data, the various confounding factors identified above cannot be quantified, because there are no data from which such a calculation can be performed, and therefore no-one can tell the extent to which they would have impacted upon the CRR for the Ultamet prosthesis. All that can be said is that they would have reduced the point estimate (and affected the confidence intervals), to an unknown and incalculable extent. Some factors would have had a greater impact than others.
- iv) It may be that (despite the evidence in SHAR) age would not have made a statistically significant difference to the data in the NJR. However, the same cannot be assumed to be true of other factors.
- v) I have no doubt that the panic engendered by the media reports in the wake of the MHRA guidance and the withdrawal of the ASR products contributed to the increased revision rates in and after 2010 to a significant, but immeasurable, extent which impacts negatively on the reliability of the data and the CRR relied on by the claimants.
- vi) The enhanced surveillance regime and the failure by some surgeons to follow the MHRA guidelines, led to more MoM hips being revised within 10 years than would have been revised had they been monitored in the same way as other articulations. The asymmetric monitoring of MoP hips and MoM hips

also affects the comparison because if the monitoring had been the same, some of the former group would have been revised within 10 years, as osteolysis would have been picked up on X-ray before it became symptomatic. It is impossible to quantify how this would have affected the CRR calculations for the Pinnacle Ultamet or the comparator.

- vii) If the extra revisions of MoM hips from 2010 onwards were clinically justified, no complaint can be made about treating them as a proxy for failure. However, the panic fuelled by sensationalist media reporting had an impact on the revision rates. An unknown number of the extra revisions were carried out on patients who were mis-diagnosed with ARMD, or as a pure precaution in case the patient developed it. Others were carried out sooner than they would have been if the MHRA guidance had been adhered to, and the patients' hips would not have failed within 10 years if they had been monitored – Mrs Blake's case is a good illustration of this. These matters adversely affected the reliability of the NJR data on Pinnacle Ultamet and thus the CRR to an unknown extent, which cannot be regarded as *de minimis*.
- viii) I am not prepared to accept that the statistics would hardly change even if as many as 50% of the revisions carried out in and after 2010 were assumed to be unnecessary or premature. That was not the subject of any expert evidence. However, even if that proposition were correct, it would just serve to underline my concern as to the propriety of treating a statistical calculation of the probable incidence of revision as a proxy for implant failure, in circumstances in which the incidence of revision on which that risk has been assessed has been increased to an unknown extent for reasons unconnected with the way in which the prosthesis has actually performed in all the patients undergoing revision.

#### **4.5 THE NICE GUIDANCE**

456. The claimants contended that the 10-year CRR for Pinnacle Ultamet on the NJR data exceeded guidelines set by NICE in 2000 and 2014. Whilst they did not contend that this by itself equated to a finding of defectiveness, they submitted that these were relevant circumstances to be taken into account in evaluating the entitled expectation of safety. I accept that the 2000 guidelines form an important part of the background circumstances to which the Court is obliged to have regard.

457. In April 2000, NICE published guidance on the selection of prostheses for primary total hip replacement by the NHS. This stated as follows:

*“Using the most recent available evidence of clinical effectiveness, the best prostheses (using long term viability as the determinant) demonstrate a revision rate (the rate at which they need to be replaced) of 10% or less at 10 years. This should be regarded as the current “benchmark” in the selection of prostheses for primary Total Hip Replacement”.*

The 10% revision rate at 10 years was based upon, and broadly consistent with the data in the 2000 SHAR report. Although it was higher than the reported CRR over 10

years for cemented implants (which was nearer 5%) and for all implants irrespective of fixation, it was much lower than the rates recorded in that report for uncemented hips.

458. The NICE guidance in 2000 was necessarily an overview, and it expressly recognised the limitations in the data that was then available from SHAR. That was why NICE recommended the establishment of a National Joint Registry, which came about three years later.
459. It is important to bear in mind that the “benchmark” set by NICE in 2000 was not a test of safety, but a yardstick for selection of the *most suitable* prostheses for implantation, based on their anticipated survivorship at that time. It was a target and not, as the claimants suggested, an indication of the maximum revision rate that a hip implant should not exceed for the purposes of use within the NHS. When that guidance was issued, surgeons only had the SHAR data with which to evaluate the likely future survivorship of a hip implant at the time of the primary operation. Although the benchmark is one of performance rather than safety, it necessarily accepts that at the time when the Pinnacle Ultamet prosthesis went onto the market, a small but not insignificant percentage of even the best performing hip prostheses could not have been expected to last for 10 years.
460. Moreover, the figure of a CRR of 10% at 10 years took no account of the higher failure rates for implants in younger and more active patients reported in SHAR 2000. The benchmark, as Professor MacGregor accepted, made no distinction between the different types of patient, such as the more active or younger patient and the elderly, inactive patients (who were far more likely to receive implants at that time) but provided a yardstick across the board. NICE and orthopaedic surgeons generally would have known that the benchmark was unlikely to have been achieved in practice if an existing uncemented metal on conventional polyethylene prosthesis was implanted in a younger or more active patient.
461. The claimants were right to avoid the suggestion that the NICE guidelines were a suitable basis for determining the entitled expectation of safety for the purposes of s.3 of the Act. The relevance of those guidelines is that in 2000, and in 2002, a hip prosthesis which had a CRR of 10% at 10 years would not have been regarded by anyone as falling below an acceptable standard of performance, let alone an acceptable standard of *safety*, irrespective of whether some prostheses were expected to do better or actually had better survivorship rates over 10 years, particularly if it was an implant that was targeted at younger and more active patients. The true CRR of the Pinnacle Ultamet prosthesis over 10 years could well be in the region of 10% or better, given all the confounding factors affecting the NJR data.
462. It does not follow that a hip prosthesis that turned out to have a higher failure rate at 10 years than the target set by NICE fell below the entitled expectation of safety, even if the aim for the future was that clinicians working in the NHS would select for use the best prostheses, i.e. those prostheses of which no more than 10% were believed to be likely to fail within that timeframe. All other factors being equal, such a prosthesis might be regarded as less suitable for use, or less clinically effective, than a product which met the target revision rate; but all other factors might not be equal, and such factors as the ease of revision and the risk of dislocation might have an impact on that assessment, as NICE itself acknowledged in the guidance.

463. In any event, the fact that there are or may be better products on the market used for the same purpose does not in itself make the product that is the “worst in class” unsafe. Thus if, for example, 12% or 13% of prosthesis type A were likely to require revision within 10 years, it does not necessarily follow from the fact that that prosthesis did not meet the target in the NICE guidelines that it would not meet the expectation of performance (measured solely in terms of survivorship) let alone the expectation of safety, that the public generally was entitled to expect. That would remain the position even if it were demonstrated that a different type of prosthesis, prosthesis B, had a CRR of only 6% within 10 years, i.e. its performance was twice as good as that of prosthesis A. Those facts might well lead to prosthesis A being withdrawn from the market over time, because it was comprehensively out-performed by prosthesis B, which clinicians who became aware of the revision rates would plainly choose in preference; but it would not necessarily follow that prosthesis A was defective.
464. On the other hand, if 50% of prosthesis A required revision within 10 years because of some common characteristic that caused them to fail, and at least 90% of all other types of prosthesis that were available at the time that A came on the market would last for 10 years or more, one could more easily reach the view that, irrespective of any other beneficial features it might have, prosthesis A not only failed to meet performance expectations but also fell below the standard of safety that the public generally was entitled to expect.
465. This illustrates that when one is measuring safety in terms of the comparative risk of revision, the question whether an implant meets the objective standard of safety set out in s.3 of the Act must be one of fact and degree. That is one reason why the Court must approach statistics relating to the actual or predicted incidence of failure of different implants over time with considerable care and caution, even if it can be shown that the data on which those statistics are based are robust and reliable. In this case, for the reasons I have already stated, the data are insufficiently reliable.
466. The benchmark was revised downwards by NICE to 5% or less at 10 years in 2014, again on an across the board approach to patients. This was presumably a response to the improved performance of prostheses generally demonstrated in the SHAR 2014 Report, most pertinently for the latest period from 2005-2014 when HXLPE had gradually overtaken MoP as the most popular articulating surface. NICE may also have relied on the data emerging from the NJR, though the latter still had less than 10 years’ worth of reliable data at that time. By 2014, the Ultamet was no longer on the market.
467. Plainly, no expectation as to the survivorship of a hip implant could have been formed in 2002 on the basis of guidance produced 12 years later. Professor MacGregor said in cross-examination that there was clearly no possible way that a 2014 standard could be applied earlier than 2014. That evidence was somewhat at odds with his written report, which suggested that the NICE 2014 guidance reflected the “minimum acceptable level of performance” of hip prostheses from 2000 onwards. In his oral evidence Professor MacGregor explained that what he meant by this was that by 2014 NICE had sufficient data to impose the new, lower, benchmark. However, he added: *“in terms of what the expectation might be in that period for a patient receiving an implant in comparison with others, that is the level at which one would have expected an implant to perform”*.

468. It was then put to him by Mr Spencer that a person could only have that expectation once the data had accumulated and they were in roughly the same position as NICE were in 2014, and he agreed. If all that Professor MacGregor was suggesting was that from 2014 onwards, armed with that information, a person might be entitled to expect a hip prosthesis to meet the 2014 benchmark, even if the primary surgery took place before 2014, that is a respectable point of view, but it does not really assist me in evaluating what the public would have been entitled to expect in 2002. If and insofar as he was suggesting that the 2014 guidance should have a bearing on determining the entitled expectation of safety at the time the Pinnacle Ultamet went onto the market, I respectfully disagree. In any event the circumstances that are relevant to that evaluation are not a matter for expert evidence.
469. The claimants submitted that the fact that the data was not drawn together and analysed by NICE until 2014 does not matter, and that the Court should take into consideration as a “relevant circumstance” the failure of the Ultamet to meet the 2014 NICE guideline. Logically that would mean that if it later transpired that the information on which a target rate for early failure of the “best performing” prosthesis was based at the time the product goes onto the market was unduly pessimistic, and the target was revised accordingly years later, the fact that the product failed to meet the revised target rate could be relevant in determining the entitled expectation of safety *even if* it met the original target. The same argument would apply by analogy to compliance with safety standards or regulations which are subsequently tightened up in the light of after-acquired knowledge, or developments in scientific research.
470. I cannot accept that proposition, which is contrary to the proper interpretation of the Act and the Directive. If the product failed to meet the original target rate by a significant margin, that might well be a relevant circumstance, but then the adjusted rate, which was lower, would contribute nothing to the evaluation. However, if the product met or was within reasonable parameters of the target rate that was thought to be acceptable at the relevant time, it does not matter whether it would have met a lower target rate which might or would have been set if NICE had known in 2000 what it knew in 2014. The 2014 guidance is of no relevance to the entitled expectation of safety.

#### **4.6 THE ENGINEERING ISSUES**

471. The claimants contended that there were various aspects of the design and manufacture of the product that increased the potential for metal wear. I have already alluded to the fact that the claimants’ materials engineering expert Professor Bull’s criticisms of the clearances in the Pinnacle MoM articulation were based on a mistaken factual premise, and both he and the claimants withdrew them. I was not greatly impressed by Professor Bull as a witness, particularly since he did not make the concession until he was cross-examined, but in the light of that concession it is unnecessary for me to address his expert report or his oral evidence any further.
472. The surviving areas of criticism included the geometry of large head MoM total hip replacements, which it was said gave rise to an increased potential for fretting, micromotion and corrosion at the modular interfaces; the roughness of the female taper of the Articul/eze head; the trunnion grooves in the Corail stem; and the geometry of that stem.

473. There is no need for me to repeat each criticism made by the claimants and the answers to it given by DePuy. DePuy's detailed response to these criticisms appears in Appendix C to their closing written submissions, which sets out the evidence of their experts, Professor Fisher and Professor Hutchings, on which they rely. I accept those submissions in their entirety; DePuy had far the better of the arguments on the engineering issues, and the advantage of being supported by two impressive experts who were plainly doing their best to help the Court, irrespective of whether their evidence was helpful or unhelpful to DePuy. In any areas of dispute on these topics, I unhesitatingly prefer the evidence of Professor Fisher and Professor Hutchings to that of the claimants' experts.
474. The claimants' criticisms were largely propounded by Professor Gill, who had been asked in his instructions to concentrate on features that potentially increased the risk of metal wear, because those were the circumstances on which his clients relied. Whilst he cannot be criticised for doing what he was instructed to do, the outcome was a report that was far less balanced than those of the defence engineering experts.
475. Professor Gill's evidence generally failed to make any allowance for the fact that in every product, an improvement in one aspect of the design may give rise to a negative feature in another aspect, but when evaluating performance and safety, all the various features must be taken into consideration, because taken in the round, the positive features may well outweigh the negatives or vice versa, or they may balance each other out. The question whether a theoretically elevated risk of wear in a specific area or component makes any material difference to the potential of the product to shed more metal debris, let alone impacts on its overall safety, must be a matter of fact and degree.
476. In this case, it was agreed by all the engineering experts that 36mm femoral heads generally (not just within the Pinnacle system) had a number of positive features. They increase a patient's range of motion (though the precise range will be influenced by product, patient and surgical factors) and they reduce the risk of dislocation. Subject to the same variable factors, they may reduce the risk of impingement. They also improve lubrication and produce less wear under *in vitro* conditions for MoM articulations (though not for MoP). A larger head typically results in a mild increase in frictional torque around the centre of the head for MoM prostheses, but this is offset by the improvement in lubrication. Decreasing clearances also result in improved lubrication. Professor Gill did not embark on any evaluation of the interplay between these features.
477. I found Professor Gill's evidence unsatisfactory. There were various respects in which his evidence turned out to be unsupported by the documents that he cited, and he was either forced to make concessions when he was challenged about them or made excuses that turned out to be unjustifiable.
478. To give but one example, he gave evidence that the ISO standard *recommended* the use of RpK as a measure of surface roughness, when in fact the ISO standard only mentions it as one possible measure and defines how it is to be calculated. It makes no recommendations as to when it should be used. When this was put to him in cross-examination, he suggested that the standard was out of date and that he had been relying on a more recent version, but that turned out to be untrue, as he was looking at the most up to date standard. In fact, at the time these hip prostheses were made, and

as Professor Gill had to concede, Ra was the standard specification used by manufacturers of hip prostheses and there was no evidence that any manufacturers used Rpk then, or even used it at the time of trial. Professor Gill's evidence, when taken at face value, was seriously misleading.

479. His evidence about the impact of larger metal heads on frictional torque was oversimplistic, concentrating on a single element in the overall complex calculation. It was easily demonstrated by DePuy that when one considered all the other forces acting on the hip joint in various directions, any increase in torque would be negligible and of no practical consequence. I regret to say that this was the point at which I ceased to have any confidence in him as a reliable independent expert.
480. I was not impressed, either, by the fact that when addressing the theory that larger MoM heads result in a lower taper connection strength, Professor Gill relied on the findings of one study by Professor MacLeod and others in 2016, to which he was a contributing author, without pointing out that the products that were tested were not Pinnacle heads, and that the engagement length of the taper used in the tests was around 5½ mm longer than that of the Corail Articul/eze mini taper ("AMT"), which would potentially make a significant difference to the outcome. There was nothing in his report that indicated that the differences in properties between the different products might mean that no reliable conclusions can be drawn from the MacLeod paper about the taper connection strength of a Pinnacle 36mm MoM head with a Corail AMT impacted using a normal surgical technique.
481. There is no standard level of roughness of a taper agreed across the industry and Professor Gill did not identify what he considered to be an appropriate level of roughness for the trunnion, or even a reasonable range of parameters. In the light of this, Professor Gill's criticism that the Corail taper has "*particularly high surface roughness generally*" is meaningless. In any event, as Professor Fisher explained, there are specific (and valid) reasons for making the interface surfaces rough, so that the microscopic form and asperities of the corresponding surfaces deform on assembly, interlock and adhere to each other.
482. The Pinnacle was a modular system, and the surgeon could use either a metal or a ceramic femoral head. This had practical advantages as well as reducing costs. The spiral groove in the Corail stem was agreed to be a design feature that reduced the risk of ceramic heads "bursting", an outcome with very serious consequences. Even with a metal head, it spreads the load and reduces micromotion. Professor Hutchings explained why fluid ingress was less likely to be a problem with a spiral groove than it is with a smooth surface. It was his view, based upon a review of the literature on this topic, that taper grooving does not lead to enhanced material loss at the head taper and may even be beneficial. I found his opinion and the reasons that he gave for it very persuasive.
483. At the end of the day, despite Mr Preston QC's valiant attempts to salvage something from the wreckage, the claimants' surviving points on the engineering issues amounted to little more than speculative theories that fell a long way short of establishing that any of the features criticised by Professor Gill gave rise to any significant increase in the likelihood of metal wear in this particular prosthesis, or the generation of higher levels of metallic debris, even if the criticisms were justified (which they were not). My impression of the Pinnacle system, taken from the



evidence as a whole, was that it was a well-designed product with many positive engineering features.

#### **4.7 WAS THE PRODUCT DEFECTIVE?**

484. The Pinnacle prosthesis using a 36mm metal femoral head articulating against the Ultamet liner was a product developed specifically to try and address the high rates of wear that promoted failure over the longer term of metal on conventional polyethylene prostheses implanted in younger and more active patients. They were the group of users for whom the product was primarily intended, although there were no indications that its use should be confined to them.
485. Those high rates of failure had been the subject of discussion over many years in the orthopaedic community, they were ascribed to osteolysis, which typically starts to develop between 7 and 10 years after implantation, but may start earlier in an active patient. It will not be noticed, however, unless the patient has an X-ray or develops symptoms. The practical experience of the orthopaedic community was endorsed by the findings of SHAR in its 2000 report which contained the only data on long-term survivorship of prostheses at the time. It was known that uncemented hip prostheses had a much higher incidence of failure than cemented prostheses, and it was expected that failure rates in both the short and long term would be exacerbated by higher activity levels. Age and sex of the patient also made a difference to survivorship of implants.
486. Those who had designed and manufactured the product expected it to deliver a substantial improvement over the existing MoP products in terms of long term wear rates, as well as various other benefits, and it was marketed on the basis of those expectations. Mr Whitwell accepted that it did deliver all the other benefits. However, the only benefit specifically relevant to safety was enhanced stability from the 36mm head, reducing the risk of dislocation.
487. It was known in 2002 that all prostheses produce wear debris irrespective of the materials used in the articulation, and that such debris could cause adverse immunological reactions which could lead to bone and/or soft tissue damage and failure of the prosthesis. It was also known in the orthopaedic community that the histology around the earlier generations of MoM hip implants that had failed in the past was different from the histology around MoP hip implants. The regulators and the surgical community were expressly warned by DePuy that there was a risk that patients might develop an adverse reaction to the metal debris from the implant and that this could cause soft tissue damage, failure of the implant or even possibly cancer. They were told, as was the case, that the risk of that happening was unknown and incalculable.
488. The Clinical Evaluation Systemic Literature review of 5 December 2008 provides a useful illustration of the breadth of information supplied by DePuy to the regulators. There is no evidence, and it was not suggested, that the regulators were misled or that anything material to their evaluation of the product's safety was withheld from them. A CE mark was granted for all the components in the Pinnacle system including the Ultamet, despite the unquantifiable risk of some patients developing ARMD.

489. All the components of the Pinnacle system met all relevant UK and European safety standards. However, there were no specific safety standards addressing what would be regarded as an acceptable rate of failure within 10 years, or the incidence of failure for osteolysis or soft tissue damage. The achievement of regulatory approval, whilst a positive factor, is therefore of limited assistance in the overall evaluation of the entitled expectation of safety in this case.
490. A producer of a new hip prosthesis would expect a learned intermediary to inform himself of any risks contained in the IFUs and technical monographs pertaining to that product and to give the patient sufficient information about them to obtain informed consent to the operation. If, in an individual case, the surgeon did not inform himself of the risks or discuss the risks with a patient, his failure to do so cannot have an adverse impact on the assessment of the objective safety of the product.
491. The received wisdom in 2002 was that a hip implant might be expected to last 10-15 years; that is what patients were told. However, there was also a risk that some might fail sooner; that was not quantified to patients. When the Ultamet was introduced, patients receiving it were told that they could expect their hip to last much longer than a conventional MoP prosthesis, because that is what the producers and the orthopaedic surgeons reasonably believed, based on the *in vitro* studies.
492. Based on the SHAR data reported in 2000, NICE had set a performance benchmark for the best prostheses of a 10% failure rate over the first 10 years, which remained unchanged until 2014. That benchmark does not reflect the reported higher rates of failure over that period of uncemented prostheses, or of hips implanted in the types of patient who were more likely to receive the Ultamet. It cannot be established that the product in this case would have failed to meet the benchmark, given the unreliability of the NJR data as a basis for assessing its likely failure rates, but even if it did, the difference was unlikely to have been statistically significant.
493. Whilst actual expectations do not necessarily correlate with what the public was *entitled* to expect, I consider that when all the relevant circumstances are taken into account, there is no justification for setting the bar higher than the actual expectations regarding short-term survivorship of existing prostheses which informed the behaviour of the orthopaedic community, prosthesis designers and manufacturers and regulators alike, and conditioned the advice that would have been given to patients in 2002, even if it transpired that everyone was mistaken and the 10-year CRR of a comparator prosthesis that would have been implanted in 2002 were in fact far better than the SHAR 2000 and 2002 reports suggested. I make it clear that I do not accept that that premise has been established on the evidence.
494. In my judgment, the public was entitled to expect that the 36mm Pinnacle MoM prosthesis, irrespective of the stem used, would not have a much greater risk of failure in the first 10 years after implantation than the *expected* failure rate over that period for the product it was designed to improve upon, namely, an uncemented metal on conventional plastic prosthesis. At that time, the expected failure rate of such a comparator was, on a conservative basis, around 15%.
495. However, even if the correct comparison for the purpose of evaluating safety is with the actual performance of a comparator, and if the data for the period from 1994-2004 in the 2014 SHAR report were to be regarded as more reliable indicators of the

performance of an uncemented metal on conventional polyethylene prosthesis than the data in the earlier SHAR reports, there is still insufficient evidence from which to be able to conclude on the balance of probabilities that the revision rates for the 36mm head Pinnacle Ultamet prosthesis were materially worse. There are far too many confounding factors affecting the reliability of the CRR for the Ultamet prosthesis, whose impact is unknown and unquantifiable, for the Court to be able to reach that conclusion.

496. I do not accept that the NJR data and the predicted risk of revision based on those data are a sufficiently robust and reliable basis for concluding that the CRR of the Pinnacle Ultamet product over 10 years were anything like 13.98%. Even if one took that point estimate at face value, it would not be a reliable indicator of the failure rates or the safety of the Ultamet product over that period because of the confounding factors that I have identified.
497. In 2002 nobody could or would have expected such low 10-year CRR as those reported in the NJR 2017 for uncemented prostheses with metal on polyethylene articulations. Those figures are not representative of the true performance of a comparator prosthesis, because they are based on data which includes HXLPE articulations and products which were not available at the time the Ultamet came on the market. The public certainly could not and would not have expected, nor would they have been entitled to expect, such low CRR in a group comprising some cemented MoP prostheses, but a far larger number of uncemented MoP prostheses, that were implanted in a range of patients with an average age of around 66-67, most of whom were younger than 70 and many of whom were active. Their expectations would have been conditioned by the data in the SHAR 2000 report and the practical experience of the performance of MoP prostheses in the preceding decade, which informed the views of the orthopaedic community and the information that they disseminated through publications, conferences and in advice given to patients.
498. There is insufficiently reliable evidence to establish that the Ultamet did have a materially worse failure rate than either the rate that was expected of a comparator at the time, or the actual failure rate of a comparator (insofar as it is possible to make a reliable assessment of the latter).
499. Indeed, because the statistics are based on inadequate data, both in respect of the performance of the comparator and the performance of the Ultamet, the claimants cannot prove that the latter had a materially worse failure rate over the first 10 years than the actual failure rate of a comparator available at the time when it came onto the market. Even on the hypothesis, which I do not accept, that the NJR data for all MoP prostheses is a more reliable source of information about the CRR for a comparator prosthesis available on the market in 2002 than the data from SHAR, the claimants cannot establish that the Ultamet failed to meet the entitled expectation of safety.
500. I therefore conclude that the claimants have failed to establish, on the balance of probabilities, that the Ultamet was defective within the meaning of section 3 of the Act (interpreted in the light of the relevant provisions of the Directive). Therefore, to the extent that any claimant did develop ARMD in consequence of being implanted with a 36mm head Pinnacle Ultamet prosthesis, DePuy is not liable to that claimant under the Act irrespective of any development risk defence that DePuy might have wished to rely upon had the preliminary issue been answered differently.

## **CHAPTER 5: THE SIX LEAD CLAIMS**

501. As I have already indicated, I have taken the evidence in the lead claims and my findings of fact in those cases into account in forming my views of the important background circumstances, particularly regarding the impact of the MHRA guidelines and the sensationalist media coverage of issues relating to MoM hip implants, the differing approaches adopted by orthopaedic surgeons to the MHRA guidelines, and the decision to revise in a given case. Although my overall conclusion that DePuy is not liable means that none of the claims in this Group Litigation can succeed, in recognition of the importance of this case to the individual claimants and in deference to the industry of counsel, it is only fair that I should end this judgment by setting out my findings in each of the six lead claims.
502. Before doing so, I must record my concerns about the contents of the witness statements. Five of the lead claimants gave evidence orally, some via video link. It was apparent that they were all truthful witnesses who were trying their best to remember events that had taken place a long time ago. However, their evidence to the Court about the impact on them of the revision (or re-revision) surgery was often very different from the content of their witness statements, which tended to suggest considerable impairment and continuing pain. When they were asked to provide an explanation for this they were unable to do so. To give but one example, Mrs Garratt's witness statement described the effect of her "injuries" as "devastating and life changing" although the claimants' orthopaedic expert in her case, Mr Smith, agreed at trial that by 2014 there did not seem to be any significant ongoing problems, and this was borne out by her medical notes. When Mrs Garratt was asked in cross-examination what she meant by this part of her evidence, she said she did not know.

### **5.1 IAN HALEY**

503. Mr Haley was born on 17 June 1946. He had a left side Pinnacle MoM implant on 8 September 2007 and revision surgery on November 2012. The metal cup and liner were replaced by a ceramic head and a polyethylene liner; the stem was left in place. DePuy accepts that Mr Haley had ARMD and that he had revision surgery as a result. His histopathology had classic features of ALVAL and is consistent with the evidence of Professor Athanasou and Professor Nelson about the correlation between high ALVAL cases and large pseudotumours. Mr Haley's case is a good illustration of the difficulties that clinicians faced in diagnosing ARMD because he had, at most, low-level symptoms, and ultrasound failed to pick up what turned out to be a very large pseudotumour that was causing muscle and soft tissue damage.
504. Mr Haley's witness statement is dated 7 December 2016, although he told the Court that this was the product of earlier drafts which originated from information he had been providing to his solicitors since shortly after the revision surgery was carried out. Like the other claimants who gave oral evidence at trial, Mr Haley very fairly accepted that he could not clearly recollect all the conversations he had had with the doctors who treated him from time to time, and that he had used the medical notes to assist his memory as well as to provide an accurate timeline in his statement.
505. In 2005, at the age of 59, Mr Haley took early retirement from his role as an operations manager. Prior to his retirement he had led an active life, enjoying many different sporting activities. He was particularly fond of golf, which he played up to

five days a week, despite experiencing some degree of pain in the area of his left hip from around 2001 onwards. He continued to play golf regularly after his retirement, but the hip pain eventually became so bad that he could only manage five or six holes before he had had enough. Mr Haley presented as the stoical type who would only go to see a doctor when the pain became unbearable, and his evidence, which I accept, was that he went for help when the pain level reached what he described as “*the chronic stage*”. That was in 2007, when he decided to seek advice from his GP, who referred him to a consultant orthopaedic surgeon, Mr Limaye.

506. Mr Limaye examined Mr Haley at Woodlands Hospital in Darlington on 6 August 2007. He noted that Mr Haley was very fit and managing most activities without problems, and that everything was normal on the right side, but that he had a long history of increasing pain and disability in the left hip. That pain was radiating down to the knee and occasionally beyond. He also had some pain in the groin, and early stiffness in the hip. X-rays showed evidence of osteoarthritis. It was Mr Limaye’s view that at the age of 61, Mr Haley was too young to have a hip replacement, and he explained to Mr Haley that a relatively young patient like himself would risk having to undergo a number of revisions over the rest of his lifetime. Despite this, he referred him for a second opinion to another consultant orthopaedic surgeon at Woodlands Hospital, Mr Eskander.
507. By the time Mr Haley saw Mr Eskander on 25 August 2007, he had given up golf altogether. The pain had limited his walking distance and was keeping him awake at night. He was also limping on the left side. X-rays confirmed that he had advanced osteoarthritis of the left hip and showed signs of early osteoarthritis of the right hip. Mr Eskander’s letter to Mr Limaye following the consultation recorded that the X-rays also indicated that there was evidence of lower lumbar degenerative disc disease.
508. Mr Eskander recommended a total left hip replacement, though he explained this would not help any radicular pain emanating from Mr Haley’s lower back. There was a short discussion about which type of prosthesis to use. Mr Eskander explained the potential benefits of a MoM hip replacement, which he said was better for fit and active younger patients, and that a MoM hip would last at least 20 years whereas a conventional cemented hip replacement may last for a shorter period. Mr Haley could not remember if anything was said about how long the conventional prosthesis would be expected to last, but he recalled Mr Eskander telling him that the Pinnacle MoM hip was “*the sort of joint we give to young fit people like yourself, we give them to lorry drivers, it will last you 20 years.*”
509. The 20 years stuck in Mr Haley’s mind because he thought that might mean that the replacement hip would last him for the rest of his life. He also remembered Mr Eskander explaining to him that the stem of the Pinnacle hip fitted into the cup without the need for “glue” (i.e. cement). Mr Haley elected for the Pinnacle hip because from what the consultant had said, he got the impression that it was a superior product that was stronger and would last for 20 years.
510. Mr Haley had the operation on 8 September 2007 and was discharged on 12 September. Prior to having the surgery, he signed a consent form which identified the risks and benefits of the procedure. Among the identified risks were dislocation, revision, and leg length discrepancy.

511. Mr Haley was “*absolutely delighted*” with his new hip. He was no longer restricted in his active pursuits. At his first post-operative orthopaedic check-up, Mr Eskander described him as progressing very well and noted that he was already walking up to 4 miles twice a day. It was safe for him to drive, swim and use an exercise bike, and he was told that he could start playing golf again in 4 weeks. At the six-month follow-up review, on 4 April 2008, Mr Haley was still progressing well. The limp had gone, he was walking speedily, and the X-rays showed no adverse features. The notes of the review state that he required long term follow-up. Mr Haley told the Court that the new hip was fantastic, and there was nothing he could not do.
512. By April 2011 a pain in the groin on the left side, which Mr Haley had been experiencing intermittently, had become sufficiently acute for him to take painkillers. He was still playing golf around three times a week. Mr Haley’s GP, Dr Smith, saw him in August 2011 and diagnosed a groin strain. She referred him for physiotherapy. An entry in the physiotherapist’s notes for 24 November 2011 recorded that the symptoms had been abolished and that Mr Haley was walking well. Mr Haley’s recollection was that the pain did not go away completely, but that the physiotherapy eased it considerably.
513. In early 2012 Mr Haley was experiencing some discomfort over the site of the scar from the arthroplasty, and he had noticed a small hernia in the muscle at the top of the wound. One evening at home he saw a TV programme in which mention was made of patients with MoM hip replacements suffering damage to their soft tissues. This prompted him to have the hernia checked out by his GP. He saw Dr Smith on 4 February and asked if he could have an MRI scan, but she explained that they did not have access to specialist (i.e. MARS) MRI equipment. She ordered an X-ray, which showed nothing of concern, but decided to refer him to Mr Eskander in case he needed checking up because of his MoM hip replacement.
514. Mr Eskander saw Mr Haley on 9 March 2012. He noted that Mr Haley had no significant problems with his left hip and that he walked long distances unimpeded. He described the small muscle hernia as being of “no clinical significance” but stated that in accordance with recent guidelines governing MoM hip replacements he was referring Mr Haley for blood measurements of his Co and Cr levels and for ultrasound examination of the left hip to confirm or rule out ALVAL lesions. If the tests were negative, he would keep him under annual radiological review, and if they were positive he would recommend referral to the revision hip centre. Mr Eskander explained that he could treat the hernia with a cortisone injection but said that he would leave it if Mr Haley was not in pain (which was the case).
515. The report of the ultrasound test failed to identify any abnormal soft tissue masses, pseudotumours or effusion in the left hip, but it did show a significant amount of fluid in the left trochanteric bursa. When Mr Haley had the next consultation with Mr Eskander, the results of the blood tests had not yet arrived. Mr Eskander told him the ultrasound findings did not suggest that there was a problem. Mr Haley was not keen to have the apparent bursitis treated, and so, subject to the results of the blood tests, Mr Eskander planned to see him again in a year’s time. Mr Haley remembered asking what would happen if the blood ion levels were high, and Mr Eskander said that in that event, Mr Haley may have to undergo a hip revision.

516. The blood test results, when they did arrive, indicated that Mr Haley's Cr levels were 11.8 ppb and Co 10.2 ppb. The following comment appears on the note of the test results: "*Levels greater than 4 mg/t (ppb) for chromium and cobalt may be associated with metallosis [local guidelines].*" This case was therefore an example of the usage of "metallosis" as a synonym for ARMD and of the use of a lower level of blood ion concentrations than the threshold in the MHRA guidance to screen MoM patients. However, Mr Haley's readings were also above the 7ppb threshold in the MHRA guidance. As he told Mr Eskander he was experiencing some discomfort in his left hip, Mr Eskander decided to refer him to the revision hip centre.

517. On 30 May 2012 Mr Haley was seen at the revision hip clinic by Mr Webb, another Consultant Orthopaedic Surgeon. Mr Webb's letter following the consultation describes the position thus:

*"[Mr Haley] seems to have done very well from his hip replacement and is very pleased with the service it gave him. Over the last year or so he has developed some aching in his groin and is naturally concerned with regards to the publicity that is around regarding metal on metal hip replacements".*

Mr Webb noted that Mr Haley was still managing to play golf although he felt he was starting to limp, and was getting some irritation if he turned over in bed at night. Mr Webb's examination of the left hip revealed mild tenderness laterally, but there was a full range of movement and no swelling. He decided to order an MRI scan to assess Mr Haley for any pseudotumours. If the MRI scan was normal and his symptoms were only low level with relatively stable metal ions "*there may be a case that we could monitor him before rushing into anything.*"

518. Mr Haley had the MRI scan in July 2012, but he did not know what the results were until he saw Mr Webb again on 10 October 2012. At that point Mr Webb broke the bad news that the MRI scan showed a large pseudotumour sitting around his hip on the left side. It was tracking distally down the thigh as well as up. There was lysis (loosening) in his proximal femur and possibly in his acetabulum as well. Mr Webb said that Mr Haley was facing revision surgery and that given the state of his MRI scan they would need to get on and do this fairly soon. He explained that the pseudotumour was around his sciatic nerve so that it would have to be dissected carefully. He told Mr Haley that removal of the pseudotumour would potentially leave him with some instability and warned of the possible risk of dislocation in the future. After discussing the pros and cons of surgery, Mr Haley agreed to have the revision, and this was carried out on 29 November 2012.

519. In the event, Mr Haley made an excellent recovery from the revision surgery. The physiotherapist who saw him two weeks after the operation reported that he was very happy with his new hip replacement, that his leg lengths were equal and that he was doing extremely well at home, with very few limitations for his stage of rehabilitation. Mr Webb saw him on 30 January 2013 and said he had done really well, despite the fact that he had a "*very nasty ALVAL type reaction with a huge amount of soft tissue destruction and [losing] his abductors and vastus lateralis as a result of the pseudotumour*". The histopathology evidence of tissue samples examined post-operatively corroborates that description. Mr Haley was walking around pain free. He was delighted with the results and was discharged from the clinic.

520. Mr Haley very fairly accepted in cross-examination that his condition by April 2013 was similar to that following primary surgery. A review by Mr Webb at that time records that he had no pain other than the odd twinge now and again. He was walking without sticks and compensating for his lack of left hip abductors. Notes taken by a cardiologist in July 2013 record that he was able to play golf and walk some 2-3 miles without too much trouble. He had a PCI stenting procedure that month and was prescribed statins thereafter, which his GP records identify as the likely cause of a period of pain and reduced mobility he suffered in the Autumn of that year.
521. Mr Haley accepted that he had suffered from a range of problems other than with his left hip that had had a significant impact on his function. Chief among these were the degenerative changes to his lumbar spine and osteoarthritis in his right hip. The expert orthopaedic surgeons in his case, Mr Kim and Professor Kay, agreed that the degenerative changes in his right hip are likely to play a role in his ongoing hip pain and ongoing mobility. Contemporaneous medical records bear this out. In November 2014 Mr Webb noted that he now had a Trendelenburg gait but was compensating for it fairly well. His left foot drop will impact on his mobility and it is likely that any residual weakness is secondary to the foot drop.
522. I accept Mr Kim's evidence that Mr Haley's reported deteriorating function and residual musculoskeletal symptoms are more likely to be secondary to referred pain from his lumbar spine and the osteoarthritis of his right hip than anything to do with his left hip or the revision surgery. He had spinal decompression surgery in May 2015 and a right total hip replacement on 3 January 2017, from which he appears to have made a good recovery. After the latter operation he had a fall, since when he has experienced intermittent pain in his right knee.
523. Mr Haley still has an altered gait, partly in consequence of the muscle damage caused by the pseudotumour, and partly as a result of the left foot drop which is related to the degenerative problem in his lower back. He has made a good recovery from the revision surgery, and most of the problems he has experienced since then are attributable to his co-morbidities. The facts of his case do not support the claimants' pleaded case that the outcome from revision procedures following pseudotumour formation is poor.

## **5.2 DIANE EMERY**

524. Mrs Emery was born on 22 March 1946. She had a right total hip arthroplasty using a Pinnacle MoM prosthesis on 20 May 2009, when she was 63. This was revised on 15 August 2012. The head and liner were exchanged for a ceramic head and a polyethylene liner; the cup and stem remained in place. DePuy accepts that Mrs Emery had ARMD which caused her to have the revision. Like Mr Haley, her histopathology results were typical of someone with significant ALVAL. After a period when all seemed to be going well, she started to develop symptoms which got steadily worse in the period leading up to the revision surgery.
525. Mrs Emery suffers from a number of co-morbidities besides osteoarthritis, including, in particular, problems with her chest and respiratory problems due to pulmonary disease. She also has long-standing problems with her back. She retired from her job as a nurse at the age of 40. She experienced a great deal of pain in her left knee for several years before she started to feel pain in her right hip in around mid-2005. In



September 2006, her GP suspected osteoarthritis and sent her for an X-ray, which showed degenerative changes present in the hips, more evident on the right side. However, the pain in her knee was far worse than the pain in her hip, and in January 2008 she had a total knee replacement, from which she appears to have made a good recovery.

526. In June 2008 she went to see her GP complaining of pain in her right hip, especially when walking. A further X-ray indicated that the osteoarthritis had progressed since her previous X-ray in September 2006 and that there was a marked loss of cartilage on the right side. Her GP therefore referred her to an orthopaedic consultant, Mr White, for advice. It appears from his letter that Mrs Emery was keen at that time to explore the possibility of resurfacing rather than a total hip arthroplasty, although she cannot now remember this.
527. Mr White examined Mrs Emery on 4 February 2009 and noted that she was struggling to mobilise because of her painful right hip. He considered that she would be a better candidate for a total hip arthroplasty than resurfacing. Mrs Emery could not recollect the discussion she had with Mr White about the pros and cons of surgery, but the consent form that she signed on 29 April 2009 identified the intended benefits as “pain relief, improve mobility, correct any deformity” and the serious or frequently occurring risks as including “infection, pain, wear, revision, dislocation and leg length discrepancy.”
528. Mrs Emery had the operation on 20 May 2009 at Llandough Hospital. Mr White used a Corail stem. After the surgery, Mr White came to see Mrs Emery on the ward and she remembered him saying to her: *“I have put a young person’s hip in that will last much longer”* to which she replied *“Oh, that’s good.”* Mrs Emery recovered well. Two weeks after the operation she was walking without crutches. Seven weeks after primary surgery she was noted to be making excellent progress. She was able to return to driving after 6 months. 8 months after the operation she was able to take long walks without using a stick, and was going to return to dancing soon.
529. In early 2011 Mrs Emery began to experience pain in her right hip which was mainly localised to the groin area and the side of her buttock. It gradually got worse, and on 25 March 2011 she went to see Mr Rath, another member of the orthopaedic team at Llandough hospital. He noted that on examination she had a good range of movement and that her X-rays were satisfactory but decided that he would order tests of her blood ion levels and, if they were raised, an MRI scan. When the results of the blood tests arrived three days later, the Cr levels (3.60 ppb) were within the range regarded by the MHRA as giving no cause for concern, but the Co levels (8.41ppb) were above that range. Mr Rath decided to order an MRI scan.
530. The results of the MRI scan in August 2011 showed that there was a thick-walled collection around the hip which tracked posteriorly around the greater trochanter. These features suggested a pseudotumour and were consistent with ARMD. There was complete fatty atrophy of the gluteus minimus, but the muscle quality was otherwise good. On 19 August 2011 Mr White wrote to Mrs Emery to explain the results of the MRI scan. He said that two of his colleagues specialised in dealing with this type of case, and transferred her to Mr Jones, a senior lecturer in orthopaedics and a consultant orthopaedic surgeon, who is a well-known surgeon in this field.

531. Mrs Emery first saw Mr Jones on 19 September 2011. Mr Jones noted that for two years Mrs Emery had been very happy with the hip, but that over the last six months matters had deteriorated, in particular with regards to pain, and that she had also developed a limp. The location of the pain was in the groin and extended around the lateral aspect of the hip. He had no hesitation in recommending an urgent revision for her. Mrs Emery recalled Mr Jones telling her that it was the MoM hip replacement that was causing the metal ions in her blood and that it needed to be urgently revised. This was the first time she had heard anything about MoM, and the last thing that she had expected to hear was that she needed to have another operation. Mr Jones' report to Mr White following the consultation stated that, depending on the amount of soft tissue damage, it was possible that Mrs Emery could have ongoing symptoms.
532. Mrs Emery underwent various further tests in November and December 2011, including a repeat blood ion test which showed that the Co levels had increased to 12.34 ppb. On 30 April 2012 she saw Mr Jones again and signed a consent form for the revision surgery. Mr Jones reported to her GP that: *"unfortunately she has significant symptoms, elevated metal ion levels and cross-sectional imaging which has demonstrated an adverse reaction to this implant. I have advised revision surgery. She understands this is hopefully to treat the symptoms that she has now but also to prevent any ongoing damage"*.
533. Mrs Emery was booked in for the revision surgery in May 2012, but she then developed a chest infection, and consequently the operation had to be postponed until August 2012. Mr Jones' notes of the operation record that (consistently with the results of the MRI scan) he found a large bulbous pseudotumour extending over the lateral and anterior aspect of the abductors; however, the underlying muscle was intact. He found a 2x1cm rent in the posterior capsule at the neck of the pseudotumour. There was extensive intracapsular necrosis and the anterior capsule was destroyed. All the affected tissue was excised. Both the cup and stem were well fixed and orientated. After the operation, Mr Jones explained to Mrs Emery that he was able just to remove and replace the head and the liner.
534. The histopathology of samples of the tissue extracted from Mrs Emery's hip during surgery showed classic signs of ALVAL. There was extensive necrosis of the soft tissue, and chronic inflammatory infiltrate with accumulations of finely granular pigmented macrophages, and focal collections of lymphocytes (perivascular cuffing). There are germinal centres and granuloma, suggesting the progression of an adaptive immune response, and metal particles can be seen clearly within the macrophages, as well as signs of corrosion products. The histology report taken at the time stated that the appearance was "suggestive of a reaction to metal."
535. Mrs Emery's explants revealed much higher wear than those of the other lead claimants. Professor Fisher's evidence was that the wear was consistent with the taper not being well fixed or assembled, and Professor Hutchings commented that the wear rate of the taper was about twice the wear rate of the bearing, which was most unusual.
536. On 4 October 2012 Mr Jones reported to Mrs Emery's GP that her progress since her revision surgery had been satisfactory, that her wound had healed without complication and that she was using one walking stick. However, he noted that the total knee replacement had become slightly more troublesome in recent weeks and

said this would be kept under review. In January 2013 Mr Jones examined Mrs Emery again, and reported to her GP that she had come through the hip revision surgery very well and that she was happy with the outcome; indeed, her Oxford hip score gave her a score of 44 out of a possible 48. He also reported that her Co and Cr levels had been checked and showed a significant decline from pre-operative levels. The results of the blood ion tests corroborate that assessment – her metal ion levels had fallen to 1.03 ppb (Co) and 0.89ppb (Cr).

537. The claimants' orthopaedic expert in Mrs Emery's case, Mr Evert Smith, suggested the Oxford hip score was akin to the sort you would see after an excellent primary total hip replacement. He thought it was so good that it must have been mis-reported, but there is nothing to support that suggestion. Mr Jones himself referred to the Oxford hip score in his letter of 21 January 2013, which is unlikely if there were any doubt about the result.
538. Although Mrs Emery subsequently experienced respiratory difficulties which led to her going to pulmonary rehabilitation, and she also experienced severe pain radiating from her neck, there were no further problems reported with her hip until 5 May 2016, when she saw her GP complaining of acute pain over the last few days, felt mostly in the right groin. The GP considered this to be mechanical pain from her lumbar spine, an opinion shared by Mr Owen, DePuy's orthopaedic expert in this case, who was not cross-examined about this. An X-ray taken in December 1998, at a time when she was experiencing lower back pain, had showed degenerative changes in her spine, and her lumbar spondylosis had got worse.
539. In September 2016 Mrs Emery asked her GP to refer her to Mr Jones to have a check on her blood ion levels, because a new consultant in the chest clinic had told her that they were still very high. Mr Jones wrote back to the GP to reassure Mrs Emery that this was not the case. In the event, she did not have a further consultation with him, and no further blood ion tests were taken. Unfortunately for Mrs Emery she had further respiratory problems which led to her spending Christmas 2016 in hospital. On discharge she was assessed as independently mobile and reported as having no regular pain on movement.
540. Mrs Emery told the Court that she still has pain in her groin and round about her hip "off and on", and that it sometimes keeps her awake at night. That is very different from the constant pain since the revision operation that she described in her witness statement. I prefer her oral evidence, which is more in keeping with the lack of any reference in her medical records to hip symptoms since the revision surgery, and the fact that she decided not to attend a follow-up consultation when afforded the opportunity to do so in April 2013.
541. The evidence in support of the conclusion that Mrs Emery suffered from ARMD necessitating revision surgery is overwhelming. The pattern of her symptoms was consistent with such a diagnosis; she developed pain around two years after her primary operation, and it continued and gradually increased in intensity up to the date of the revision. There was a large thick-walled pseudotumour, and the radiologist who carried out the MRI scan reported that the findings were compatible with a reaction to metal. The histology was entirely consistent with those findings. Overall, she is a good illustration of a patient whose symptoms and investigations pointed to a clinical diagnosis of ARMD, which was then borne out by the findings of a very experienced

specialist orthopaedic surgeon at surgery and corroborated by the histological evidence.

542. Mrs Emery appears to have made a good recovery following her revision surgery. I accept the evidence of Mr Owen in this regard, supported by his findings upon examination, which ultimately was supported by the evidence of Mr Smith. There is no evidence of any significant symptoms or impairment to her function or mobility consequential on the revision and the fact she declined to see an orthopaedic consultant when afforded the opportunity to do so in April 2013 reinforces the impression that she was doing well. Medical records relating to her other problems are consistent with this; in May 2013 a consultant chest physician, Dr Lau, recorded that Mrs Emery “*successfully underwent revision of her hip*” and an anaesthetist’s assessment of her in September 2015 prior to her epigastric hernia repair recorded that apart from her chest condition “*she has no other medical problems*”. The acute pain reported in May 2016 was of recent origin and she had not been to see a GP complaining of hip pain at any point in the intervening period. Subsequent to that episode, in December 2016, she was described as “*fully mobile*”.
543. Overall, Mrs Emery’s medical records are inconsistent with any suggestion of chronic ongoing hip pain since the revision surgery. Mr Smith accepted this and said it was a matter for the Court to decide how to resolve the discrepancies. He was a generally fair and balanced witness who volunteered in the witness box that he wished to reappraise some aspects of the evidence in his expert’s report in the light of the oral evidence given at trial. DePuy criticised his reluctance to accept that if Dr Wilson’s and Professor Nelson’s evidence were accepted in the cases of Mrs Garratt and Mrs Blake, neither of those claimants had ARMD. I consider that this was more in keeping with a degree of diffidence about trespassing into an area that was a matter for the Court to determine, than with any conscious or subconscious desire to promote the claimants’ case.
544. In the case of Mrs Emery, Mr Smith very frankly stated that he thought she had recovered satisfactorily from her surgery, he just had his doubts about the high Oxford hip score. He acknowledged that the discrepancy between her symptoms of pain as reported to him and in her witness statement, and what was stated in the medical records, was for the Court to resolve. He accepted that “*on balance, if something comes up three and a half years after the event, one might investigate, but the chances are it has got nothing to do with revision.*” When he was asked whether he thought the sudden onset of pain after 3½ years could be related to ARMD and the revision, he said he did not. He did not suggest that Mrs Emery’s recovery was any worse than that of a patient who had been revised for reasons other than ARMD.
545. At the end of the day there was a large measure of agreement between Mr Smith and Mr Owen about Mrs Emery’s case and about her good recovery from the revision operation. If she is suffering any pain around the hip at present, I find on the balance of probabilities that it is more likely to be due to her co-morbidities than anything to do with her hip revision.

### **5.3 DAWN BLAKE**

546. Mrs Blake was born on 12 August 1959. She received two Pinnacle MoM hip replacements, but her claim relates only to the left prosthesis, which was implanted on

6 December 2005 when she was only 46. That was her second total hip arthroplasty. She had received a right Pinnacle MoM prosthesis on 26 July 2005. The left prosthesis was revised on 19 March 2009. The stem, head and liner were all exchanged, but the cup remained. Her right prosthesis is still in place.

547. Mrs Blake has a particularly complex medical history. She retired from her role as a nurse in 1996 after suffering a back injury when she was attacked by a patient, but she also has severe underlying generalised osteoarthritis which has caused significant degenerative changes to her spine and problems in all her joints. She had to undergo a lumbar discectomy in 1996 and later, in 2000, she required a spinal fusion. She has long term acute back pain, not only in the lumbar region but also in the thoracic and lower cervical spine. She has also undergone surgery for carpal tunnel syndrome. She suffers from fibromyalgia, hypothyroidism and hypertension and is diabetic.
548. The first mention of hip pain in Mrs Blake's medical records was on 4 February 2005, when she told her GP she had suffered six months of right hip pain and stiffness, for which she was taking strong analgesics. At that time, she was also suffering pain in her right knee, which had been injected, though that only provided pain relief for a week. Her GP referred her to a consultant orthopaedic surgeon, Mr Taylor, at Lourdes Hospital, whose letter to the GP following that consultation on 2 June 2005 describes her as having discomfort in the whole of her knee and also particularly in her groin. The pain was significantly affecting her walking distance. X-rays confirmed that she had osteoarthritis of both hips and Mr Taylor felt that they were her main problem. His view was that although she was young (being then only 45) she needed bilateral hip replacements. Mrs Blake remembered that she was told by Mr Taylor that because of her age, if she had a hip replacement, she would have to have several revisions over the rest of her life.
549. Mr Taylor no longer performed hip surgery, so he referred Mrs Blake to a colleague, Mr Miller. He shared Mr Taylor's view that she required bilateral total hip replacements and because her right side was the more painful, he arranged to perform this operation first. She had a Pinnacle MoM hip with an S-ROM stem. On post-operative review, twelve weeks later, on 27 October 2005, Mr Miller told the GP "*I am delighted that her hip replacement is absolutely asymptomatic and looks good both clinically and radiologically.*" Unfortunately, the right knee was giving Mrs Blake increasing problems, and he arranged an MRI scan. This showed degenerative change and a small effusion.
550. On 29 November 2005 Mrs Blake signed the consent form prior to undergoing surgery on her left hip. The intended benefits were identified as pain relief and improved mobility. The serious or frequently occurring risks identified included infection, dislocation, nerve/blood vessel damage and leg length inequality. The operation was carried out on 6 December 2005, using exactly the same type of stem and femoral head as on the right side. An X-ray taken the following day showed a satisfactory appearance of the left hip prosthesis. Six weeks later, Mr Miller wrote to Mrs Blake's GP to say that he was delighted to say that all was well, commenting that this had been an easier replacement than on the right side and she had recovered faster. He would see her again in six months' time.
551. On 12 July 2006 Mr Miller examined Mrs Blake and reported to her GP that both clinically and radiologically both of her hips were extremely good, but she was still

complaining of pain in her right knee. She had begged him for a total knee replacement, which he felt loathe to carry out without confirmatory evidence, so he had persuaded her to undergo an arthroscopy. That was carried out on 18 July 2006, and she also had a lateral release. Mr Miller noted quite marked retropatellar degenerative change, although the major joint was in good condition.

552. Following her next consultation with Mr Miller on 7 December 2006, he reported to the GP that she had “*a very good result from both her hip replacements*” but unfortunately the right knee was continuing to give her the most unpleasant pain. He had agreed to arrange an MRI scan of her thoracic spine, which had also been causing her considerable pain at that time. When the MRI results were obtained, Mr Miller advised Mrs Blake that she had long standing wear and tear causing disc bulging at the bottom of her neck and in the mid thoracic area. Unfortunately, he did not think there was a surgical option for this.
553. On 12 January 2007 Mr Miller reported to the GP that Mrs Blake was having so much trouble with her knee that she wished to have a total knee replacement and that he was going to arrange for this to be carried out in March. He referred her to a pain management consultant. The total knee replacement was performed on 18 March 2007. Mrs Blake developed a post-operative infection which was successfully treated with antibiotics, but otherwise appeared to have made a good recovery.
554. Towards the end of 2007 Mrs Blake was experiencing neck pain radiating to both shoulders, especially the right. Investigations showed generalised degenerative changes throughout the cervical spine, but the consultant spine and neurosurgeon did not think this required surgery. He referred her to a consultant rheumatologist, whose consultation notes described her hip and knee operations as “*very successful*” apart from the fact that the right knee tended to be unstable. The consultant rheumatologist opined that the overall picture was of lower limb inflammatory arthropathy due to ulcerative colitis.
555. In January 2008 a pelvic X-ray indicated satisfactory alignment of the hip prostheses and no evidence of any infection or loosening. Later that month Mr Taylor wrote to Mr Miller to say that for various reasons he felt that Mrs Blake was a probable candidate for revision of her knee replacement and that he was going to refer her to a Mr Davidson. He told Mr Davidson that she was keen to have something done. Mrs Blake’s evidence was that the hip replacement had been such a success that she wanted a knee replacement. Mr Davidson noted that there was medial instability in the knee and that Mrs Blake could only walk with a crutch. He went through the pros and cons of a joint replacement with her, and advised her of the higher risk of infection to which she would be exposed, but she was keen to proceed. In February 2008 she underwent a revision right total knee replacement. Up to that point, there are no specific references in her medical records to experiencing pain in her left hip, and her back pain was at a low level and manageable with painkillers. This is consistent with the account she gave to Mr Owen (that she had no real problems with her hip until December 2008) and with her oral evidence.
556. The knee surgery appeared to be successful although the left leg was an inch shorter than the right in consequence. Following the knee revision surgery Mrs Blake still had a lot of pain in her back, for which she was visiting the pain clinic, and which affected her mood and interfered with her hobbies.

557. Four months later, on examining the revised knee, Mr Davidson reported Mrs Blake as pain free and very pleased with the results, but she had been getting some left sided thigh pain and knee pain. He suspected some of the pain was coming from her hip, as the knee was normal on examination. The hip replacement looked satisfactory on X-ray and Mr Davidson suspected that once she had a shoe raise, some of her symptoms might improve. The X-ray results indicated that there was no evidence of loosening or dislocation and both hips were satisfactorily aligned. Minor degenerative changes were noted in the left knee.
558. The first mention in Mrs Blake's medical records of any specific problem with the revised left hip is in a letter to her GP on 3 February 2009 from Mr Dunlop, the consultant who took over many of Mr Miller's patients when Mr Miller retired. Mr Dunlop gave evidence at trial. He was a frank, straightforward witness. Mrs Blake told him that she felt continuing and worsening pain deep in the joint and, if anything, posterior which he commented "*seems to have crept up on her insidiously over the years since she had the initial procedure performed.*" She also told him that there were no wound problems around the time of the surgery and "*initially the surgery was a success*". In his oral evidence, Mr Dunlop interpreted his note as meaning that the pain had been gradually building up over some time, though not necessarily going all the way back to the primary surgery.
559. Mrs Blake's medical records contain one later entry in which reference is made to a long-term problem: in October 2014 she is recorded as having told another orthopaedic surgeon, Mr Middleton, that "unfortunately the left hip *never completely played ball* and it required revision by Colin Dunlop". However, that is not reflected in Mrs Blake's contemporaneous medical records in the period prior to February 2009. It is also difficult to reconcile with Mrs Blake's statement that the hip replacement had been such a success that in January 2008 she wanted her knee joint replaced. It seems more likely that the pain arose later in 2008, which is what she told Mr Owen.
560. DePuy sought to rely on the fact that Mrs Blake told Mr Owen that she had walked with a stick in her right hand since primary surgery, which he observed in his oral evidence "*would imply that she has potentially some weakness in the left hip*", explaining that people use a stick in the opposite hand for support. Although I accept what Mr Owen says about that, in the light of Mrs Blake's other problems, particularly the pain in her back, it would be dangerous to make an assumption that she was using the stick because her left hip was painful.
561. Mrs Blake told the Court that her recollection was that after the knee was revised she was happy with everything until later in 2008, when she began to get a sensation of creaking in the joint when she walked and a constant pain, like a nagging toothache. She put up with it for a while, and then decided to have it checked. She told Mr Dunlop in February 2009 that the pain was bothering her at rest as well as when she was walking. On examination she was a little tender over the trochanteric region, but she told Mr Dunlop that this wound pain was not the pain she experienced deep within the joint. On examination she was Trendelenburg negative, an indicator that she did not have a hip problem, though not a conclusive one.
562. Mr Dunlop told the Court that he would have had Mrs Blake's full medical records available to him at the time of his initial consultation with her, and that he did not

think her ongoing back problems or fibromyalgia were relevant to the problem that she was presenting. If he had thought they were relevant, he would have mentioned them in his letter to her GP. He considered there were two possible reasons for her symptoms; the first was an infection, the second, a problem associated with her MoM prosthesis, possibly an adverse reaction. He told the GP that he would review Mrs Blake again when he had all the results of the tests he had ordered to hand, and hopefully they could make a long-term plan at that stage.

563. At that stage Mr Dunlop had his own personal concerns about MoM implants. Mr Dunlop said the consensus at the time among his colleagues in the hospital and fellow members of the local forum of hip and knee surgeons was that increasing symptoms in a patient with a MoM hip prosthesis would warrant (surgical) exploration, and possibly revision. If he were to explore a MoM hip at that stage, the minimum he would have done, having opened it up, would have been to exchange the head and liner. However, before carrying out an exploration, he first wanted to rule out any other possible causes of Mrs Blake's symptoms.
564. Mr Dunlop arranged certain investigations, including an ultrasound scan to look for any fluid collection around the hip which might be indicative of an allergic or vascular-type problem (i.e. ARMD). A radiological report largely ruled out the possibility of any prosthetic infection. X-rays of the pelvis suggested that there might be some loosening of the femoral component on the left, but Mr Dunlop appears to have discounted that as a likely explanation, and the orthopaedic experts' view is that he would have been right to have done so.
565. The ultrasound was carried out by Dr Aniq, an experienced consultant radiologist, who was told that this was a painful left hip and "*a possible case of metal allergy*". The result of the ultrasound test is dated 9 February 2009. It stated that the left hip joint was normal and there was no evidence of any joint effusion. The iliopsoas muscle was normal. There was an 8 x 4cm fluid collection noted adjacent to the posterior facet of the greater trochanter in keeping with trochanteric bursitis. There was a 4mm area of calcification/ossification noted in the trochanteric bursa as well. There was a small amount of fluid noted in the region of the gluteus medius and gluteus minimum tendon, but no evidence of tendonosis.
566. In cross-examination Mr Dunlop appeared reluctant to accept that there was nothing in the ultrasound test results to suggest any adverse reaction to metal debris, apparently because of the fluid collection adjacent to the trochanter. Yet when he described the ultrasound findings in his letter to Mrs Blake's GP following a further consultation with Mrs Blake on 24 February 2009, he repeated Dr Aniq's findings verbatim, including that the fluid collection was in keeping with bursitis, and added "*there was no collection of fluid within the joint itself and no gross tendinosis or pseudotumour mass around the hip*". That suggests that at the time, Mr Dunlop did not regard the ultrasound as lending any support to the thesis that this was a case of ARMD. However, a concern that ARMD might be the cause of the pain was still playing on his mind. The fact that the ultrasound indicated the hip to be normal did not dissuade him from carrying out a surgical exploration.
567. At that time, the hospital in Liverpool where Mr Dunlop worked did not have a MARS scanner that would enable reliable MRI scans to be taken of a patient with a metal on metal prosthesis, so he did not order an MRI scan. Nor did he order blood



ion tests; he explained to the Court that the local laboratory did not perform such tests, and there would have been a cost involved in sending samples away. Mr Dunlop was not too sure what the “normal” range for cobalt and chromium levels in patients with MoM implants was, and in any event, he agreed when it was put to him that the results of any blood ion tests would not have altered his decision to perform the surgery. Mr Dunlop’s mindset was that unless he found a positive alternative explanation for the symptoms described to him, he saw no alternative to opening the patient up to see for himself what (if anything) was going on.

568. On 24 February, Mr Dunlop explained to Mrs Blake that the problem might simply be bursitis, but he told her GP that “*in the back of my mind the metal on metal joint and metal allergy problem is a possibility*”. As a first step, he would aspirate the joint and inject the bursa with a local anaesthetic and a steroid. He would contact her by telephone afterwards to find out if this had aided her symptoms. If it had no impact, he saw no other option but to explore the joint, replace the MoM bearing and assess the components intraoperatively for any loosening. That would also allow him to excise any bursa material and assess the abductor repair as a possible other source of her discomfort. Mrs Blake’s recollection was that Mr Dunlop said to her that he could not make a decision about what to do until he had gone in and found out what the problem was, but if he needed to revise the hip, he would revise it. However, it appears from his letter to the GP that Mr Dunlop was going to replace the metal components of the prosthesis come what may - irrespective of whether there were any signs that she was suffering from ARMD.
569. Mr Dunlop undertook an aspiration of the hip on 28 February and the joint was dry. Although a so-called “dry tap” can be the result of faulty equipment, or not putting the needle in the right place, it can also be because there is nothing to aspirate. Taking into consideration all the other clinical evidence regarding Mrs Blake’s condition at the time, it is more likely than not that there was no fluid to aspirate. Mr Dunlop did not recall having any technical difficulty with this procedure, and if he had done, I am satisfied that he would have said something about it to Mrs Blake when he told her what the results were. Mrs Blake’s recollection was that she was told that the aspiration was dry, and that Mr Dunlop was worried; he realised there was something wrong, but he could not diagnose it. Mr Dunlop accepted that that was a fair assessment of his position at the time.
570. There is no note of any further discussion between Mr Dunlop and Mrs Blake before the operation took place, but it is likely that there was a telephone conversation as planned to find out whether the injection had helped. However, it failed to alleviate the pain, and Mrs Blake underwent a pre-operative nursing assessment on 16 March. On 19 March she signed a form consenting to the “*exploration of left hip and revision of bearing and femoral component*”. The operation took place on the same day, approximately six weeks after Mr Dunlop first saw Mrs Blake in clinic.
571. Despite the fact that none of the investigations Mr Dunlop had ordered had resulted in any positive signs of ARMD, he felt at the time that the combination of a MoM hip and unexplained pain was sufficient to justify a revision. He said he would have been reluctant to postpone surgery for another three or six months and monitor Mrs Blake to see how matters developed because “*if there was an evolving problem, to sit and watch and wait with a potential toxicity problem, I was uncomfortable*.” He also said that the creaking symptom she described would have pushed him towards revision

surgery even if the femoral head had been ceramic rather than metal. He said: “*I will revise a ceramic bearing for pain and creaking if the patient wants me to.*” However, he described his approach as “*more aggressive*” where the prosthesis was MoM.

572. Mr Dunlop’s operation notes record:

“Incision through old and extended.

Metallosis noted on incising fascia lata, swab to microbiology.

Abductor repair failed – metallosis in bursa.

Thorough debridement – samples to histology & microbiology.”

Mr Dunlop agreed that by metallosis, he meant the greying or discolouration of tissue due to metal debris, and that discolouration would not, in and of itself, indicate soft tissue damage, as some discolouration would be expected in any patient with a MoM hip replacement. However, Mr Dunlop thought that he must have been using the term “metallosis” in this context to include an appearance consistent with synovitis of the soft tissues around the hip joint. For some reason he appeared to draw that inference from the size of the specimens of tissue he sent for histological examination from both the sites where metallosis was noted, though the connection is not easily apparent. He accepted that synovitis could not be diagnosed until there had been histological examination of the tissues, but he explained that the surgeon can see a florid reaction in the tissues, with proliferation in the lining of the joint, and he assumed that this is what he saw on this occasion.

573. I find that Mr Dunlop is mistaken in that regard. There was no note of any florid reaction in the surrounding tissues or use of the word “synovitis” as I find there would have been if Mr Dunlop had seen any. This was not a surgeon who had a cavalier attitude to keeping accurate records. It must be remembered that he was expecting to find signs of an adverse reaction to the metal on metal hip, and therefore it is more likely than not that he would have expressly recorded any obvious signs of such a reaction if he found them. Mr Dunlop did not record any soft tissue damage. He said, and I accept, that he would not have recorded a minor amount of damage to the soft tissues because the consequences of that were likely to be insignificant in terms of pain. In the analysis of the excised tissue by the histopathologist there was no inflammation noted, although some metal debris was discovered.

574. In his witness statement Mr Dunlop originally stated that fluid that was noted pre-operatively on the ultrasound scan was both in the bursa and communicating with the hip joint itself; but in his oral evidence he corrected this, and said he had made an assumption that the fluid was communicating with the hip joint. In cross-examination he explained that this was an inference he drew, because he noted in his surgical report that there was a failure of a repair to the abductor. When he was asked about this failure by Mr Antelme, he said this:

*“In my experience, if you have an abductor repair, failure usually occurs early on and usually the hip would never have functioned terribly well, and the pain is often present from early on. This could have been a recent failure and that could have been responsible for the recent pain, I don’t know.”*

575. Mr Dunlop also decided to revise the stem. He explained that this was because he felt that the sleeve at the top of the stem that adheres to the bone may have debonded and become loose. He said he wanted to do one good operation and not leave anything to chance. He noted in the operation notes that the collar (i.e. the sleeve) was *possibly* loose. In the event, he had great difficulty in removing the S-ROM stem, which was well fixed, and had to perform an osteotomy. He inserted a new stem and a ceramic head and liner.
576. Mrs Blake's evidence was that her husband told her that Mr Dunlop had been to see him before she came out of surgery and told him that she had nearly died during the operation. Mr Dunlop was visibly shocked by that suggestion when he gave his evidence. Although there is no doubt that the operation proved to be less straightforward than Mr Dunlop would have wished, because of the problems in removing the stem, Mr Dunlop would have remembered if matters had taken such a drastic turn, and there would have been something about it in the contemporaneous medical records. It is far more likely that this was a gloss that Mrs Blake's husband put on whatever Mr Dunlop had said to him about the difficulties he had encountered. It does appear that Mrs Blake lost a lot of blood and had to have a transfusion. I have no doubt that Mrs Blake genuinely believed that her life had been in danger, and that this conditioned her view about the operation and what led up to it.
577. Six weeks after the operation, Mr Dunlop reviewed Mrs Blake in his clinic and reported that she was making good progress and walking with two crutches. She reported that she whilst she had significant pain until 1-2 weeks ago, she had noted a significant improvement in her symptoms since then. He reassured her regarding the histological findings. In that regard, Mr Dunlop agreed with Mrs Blake's evidence that he told her that she had suffered an adverse reaction to the metal, that grey tissue had been cut out, and that all the metal debris had been removed.
578. On the final occasion when he saw her, three months later, on 16 June 2009, Mr Dunlop said that Mrs Blake was a little impatient with her total left hip replacement, but she seemed to be making satisfactory progress. She was now walking without crutches and was Trendelenberg negative on both sides, a good sign. She had some pain over the anterior thigh which recurs with mobilisation, which he suspected would settle in the fullness of time. Her X- rays were satisfactory.
579. On 10 September 2009 Mrs Blake had a primary lumbar discectomy C6/7 and fusion. She and her husband then moved to Greece, where she had further neck surgery. In 2014, having returned to England, Mrs Blake was referred by her GP to another consultant orthopaedic surgeon, Mr Middleton. She was worried because her right hip was also a Pinnacle MoM and she said in her oral evidence she was "*paranoid that she had problems on that side*" given what she had been through with her left hip. Mr Middleton's notes reflect that he understood from her description of that experience that she had suffered a "*severe metallosis reaction*". He recorded that since that time (i.e. since Mr Dunlop's revision) that hip had never completely played ball and she had needed to walk with a stick with a significant limp and discomfort over the lateral aspect of her thigh. Mrs Blake said in her oral evidence that she had never been pain-free in her left hip since it was revised, but that the pain prior to the revision was in a different place and it was unremitting.

580. In the light of the history described to him in such terms, and her apparently increasing decline in function in the right hip, Mr Middleton decided that he agreed with Mrs Blake's concern that a similar adverse reaction was occurring in her right hip, and so he organised blood ion tests and a MARS scan of the right hip. This showed significant wasting of the gluteus minimus and medius muscles, partial avulsion of the gluteus medius tendon from the greater trochanter and a small thin walled fluid collection adjacent to the greater trochanter which measured 4 x 1 x 2.5cm. The blood ion tests revealed the chromium and cobalt levels to be within limits that gave no cause for concern.
581. By the end of 2014 Mr Bartlett had taken over from Mr Middleton as the treating consultant. Mr Findlay, the member of Mr Bartlett's team who reviewed Mrs Blake on 10 December 2014, told her that the results of her scan and metal ions were encouraging, but they felt that in view of her poor Oxford hip score of 13 (with pain in both hips and severe pain in her right groin) they should aspirate her hip, send samples for metal ions and culture, and inject local anaesthetic. He noted that "*she had a rather poor outcome following her revision. She is keen that this does not happen on the right hand side.*"
582. The attempt to aspirate the right hip joint proved unsuccessful; this time the notes make that clear. In March 2015 a letter from Mr Bartlett to Mrs Blake's GP after his first consultation with her recorded that she had noted a slow deterioration in the function of her right hip with progressively intrusive groin pain, particularly with active flexion such as driving a car. By then, her Oxford hip score had gone down to 9. He commented: "*it has been a difficult diagnostic process of trying to establish the exact cause of her right groin and leg pain symptoms. The concern from the outset has been that she has had a metal on metal reaction, however the subsequent investigations we have performed in the form of a MARS scan and metal blood level ions have all pointed away from an ALVAL reaction. She also has normal ESR and CRP levels pointing away from infection and the X-ray itself shows a well-fixed implant with no evidence of lysis or loosening.*"
583. Mr Bartlett considered that the second potential source of her right hip pain was possible referred pain from the lumbar spine. He was not disposed to recommend revision surgery on the basis of the evidence they had. He said that importantly, her diagnostic hip injection was "strongly negative". He suggested a repetition of the MARS scan and metal ion blood tests, as it had been approximately six months since the last series of investigations, and if the situation around her right hip appeared to be developing a metal on metal reaction his advice might change.
584. Following those tests, the results of which were very similar to the previous findings, Mr Bartlett wrote to Mrs Blake's GP saying that although the pain was causing Mrs Blake significant disability, there was no appearance of any significant ALVAL reaction around her right hip replacement (another example of the use of "ALVAL" to mean ARMD) or any clear evidence of lysis or loosening of the implant itself. There was no targetable reason for having a revision other than unexplained pain. However, after careful consideration she had decided to put herself forward for a revision right total hip replacement. Although Mrs Blake was then listed for revision surgery on her right hip, her co-morbidities intervened. A problem with her shoulders became more acute and she had to have surgery on them. She was then referred to a consultant

rheumatologist in order to explore the possibility that the pain that she was suffering throughout her joints was due to a severe inflammatory process.

585. In March 2016 Mrs Blake wrote a letter to her GP in which she said she knew there was a discussion about whether her right hip truly needed revision or not and that she had had over a year of trying to decide what to do. She made the decision to go for revision based on her pain, the difficulty she had in driving any distance, but also and most importantly, because her metal on metal hip on the left had “*failed drastically after three years and had all but destroyed her muscle in that leg and hip*”. That description of what had happened to her left hip prosthesis bears no resemblance to the contemporaneous medical notes; whilst I have no doubt that Mrs Blake sincerely believes that the primary implantation was a disaster, it would appear that over time she has convinced herself that matters were far worse than they really were. She subsequently underwent patient assessment for the revision surgery on the right hip, but it never took place.
586. In April 2017 Mr Bartlett, who was clearly still having doubts about the advisability of revision surgery on the right hip, thought that the pain might be referred pain from her spine, and referred Mrs Blake to a specialist spinal surgeon, Mr Khan. On 29 April 2017 Mr Khan agreed that the presenting complaint was now more suggestive of spinal origin than of hip pathology. He described how Mrs Blake had pain on both sides, but it was spreading down through the thighs down to her calves and that he considered it was more likely to be [a result of] lumbar stenosis. In consequence of that, it is no longer planned or anticipated that Mrs Blake should have a right sided hip revision.
587. Taken together, Mrs Blake’s symptoms prior to the revision surgery do not point towards a diagnosis of ARMD. The Hardinge approach, which was used in the primary surgery, involves cutting through the main muscle that controls the stability of the pelvis and stitching it back afterwards. A failure of the abductor repair following the Hardinge approach is quite a common occurrence, and it can lead to trochanteric bursitis. Unlike other surgeons in the lead cases who were faced with this possibility, to his credit Mr Dunlop did at least consider it. At surgery Mr Dunlop recorded that the abductor repair had failed. That was Mr Owen’s opinion of what was likely to have happened, and Mr Smith agreed that this might have been the cause of the problems that Mrs Blake was experiencing. On the balance of probabilities that would appear to be right.
588. Whilst pain “deep within the joint” is not characteristic of trochanteric bursitis, nor typical of pain caused by a failure of the abductor muscle repair after a Hardinge procedure, in the light of Mrs Blake’s many co-morbidities and the subsequent history with her right hip, it is difficult to be sure where the pain she described to Mr Dunlop was coming from. There was a dispute between Mr Smith and Mr Owen as to whether it was consistent with fibromyalgia, but in their final submissions the claimants’ counsel rightly accepted that it was potentially consistent with referred pain, which is now thought to be the explanation for her similar symptoms on the other side.
589. There is conflicting evidence as to when Mrs Blake first started to experience pain in and around the hip following the primary surgery. Given her various co-morbidities that is hardly surprising. The typical patient with ARMD will be initially asymptomatic, and will develop increasing pain over time, but there can be

exceptions: as Mr Haley's case suggests, someone can have a large pseudotumour without experiencing any significant pain at all (though the evidence in that regard could be attributable to Mr Haley's stoicism). Whilst Mr Dunlop had the impression from what Mrs Blake told him in February 2009 that the pain had crept up on her insidiously over the years, it seems more likely from the medical records to have been of more recent origin. Whilst the timing and location of the pain could be consistent with ARMD, that is not the only explanation, nor is it the most likely one.

590. The ultrasound performed on 9 February 2009 pointed against a diagnosis of ARMD but was in keeping with trochanteric bursitis. The size of the extra-articular collection, 8cm x 4cm, was within the range for a trochanteric bursitis, albeit slightly larger than normal. The expert radiologists, Dr Ostlere and Dr Wilson, accepted that there was no record of any communication between the collection observed on the imaging and the joint. They did not see any images in this case, so had to rely on the contemporaneous report. Of course, there is a possibility that Dr Aniq failed to spot a communication, but this seems highly unlikely, given that Mr Dunlop had specifically asked him to look for *"any fluid collection around the hip which may be indicative of an allergic or vascular type problem with the hip"*.
591. Dr Aniq was a consultant radiologist and a specialist musculoskeletal practitioner, and Mr Dunlop accepted that he would know what he was looking for in this type of investigation. Both radiology experts indicated that an experienced consultant radiologist was less likely to miss a connection, particularly if he was aware of the suspected pathology. They also agreed that if a consultant radiologist noted a collection around the greater trochanter he would be likely to look for the source of the collection. Dr Wilson said: *"the instant reaction of any competent radiologist would be to say where is it coming from and then to track it back, and ultrasound is a particularly effective technique at that because you can run the probe back to the ends and look and see if it's tapering down"*.
592. I was unimpressed by Dr Ostlere's argument that because there was nothing in Dr Aniq's report to indicate that the posterior aspect of the hip was examined, one could not reliably draw the inference from the report that there was no communication between the collection and the joint. As Dr Wilson explained in his evidence, radiologists do not usually record negatives. That evidence accords with sense and logic – it would be a waste of time to note down everything seen that was normal. One can safely assume that the radiologist will only note down anything on the images that he considers to be significant, and that sufficient images will be taken to enable the radiologist to form a view about the overall effect of what he has seen.
593. Dr Ostlere's approach to the radiological evidence in the lead claims appears to have been conditioned by his view, articulated in his oral evidence, that whenever there was a mass on the lateral side of the hip, one should assume it was a case of ARMD until proven otherwise. Whilst I do not doubt the sincerity of that view, it appears to have subconsciously fuelled an inclination on his part to find symptoms where none existed, and to explain away evidence that was either neutral or inconsistent with ARMD. I found some of his explanations or qualifications of Dr Wilson's evidence, such as the one I have just referred to, implausible. Dr Ostlere candidly accepted that his approach had been to look for evidence to support the claimants' case. He also appeared to be unduly swayed in his conclusions by the views of experts in other

disciplines. I found Dr Wilson's evidence more balanced and helpful, and preferred it in areas of disagreement between them.

594. The absence of any evidence that the collection communicated with the hip is consistent with the fact that the hip aspirate was dry. Of course, a "dry tap" is not conclusive evidence of the absence of fluid in the hip joint, because that result could be due to technical difficulties during the procedure. However, there was no evidence that Mr Dunlop encountered any such difficulties in this case; none was recorded, and Mr Owen said that for somebody of Mr Dunlop's experience, a dry tap would be highly unusual. Mr Owen explained that if he had been actively looking for a joint effusion and got a dry tap, he would not just give up. He would speak to his radiology colleagues and perform the procedure again. That makes sense and accords with the likely approach of any surgeon. On the balance of probabilities, the combination of the dry tap and the radiology results in Mrs Blake's case point firmly against a diagnosis of ARMD.
595. Next, I must consider the surgical findings by Mr Dunlop and the histopathology. There is no mention in the operation note of any necrosis of the soft tissues around the joint, which is the key feature of ARMD. There is no mention of any abnormal fluid within the hip joint or any hip effusion, and no mention of a pseudotumour. Whilst there was an abductor repair failure, the main muscles were intact, and whilst there was metallosis (metal staining) in the bursal tissue and the capsule, there was no obvious damage to the surrounding periarticular tissues. One cannot assume that because there was metallosis, there had to be a communication with the hip joint. Mr Dunlop did not notice any communication between the collection and the hip joint; as he clarified in his oral evidence, he simply assumed that there was one. A swab was taken, indicating that some fluid was present, but that is far too unspecific a finding for any inferences to be drawn from that – a collection of fluid was seen on the ultrasound. Mr Dunlop described the tissue he excised as "*abnormal*" and sent it off for histological examination, but I am satisfied that that was because of the metal staining and not because of anything else he had noticed. I was not persuaded by Mr Smith's evidence that the clinical findings at revision were indicative of ARMD.
596. As Mr Owen suggested, and Mr Smith accepted, the metal debris that caused the staining could have come from the junction of the collar and the stem, which is an area much closer to the bursa. Alternatively, the metal debris could have reached the bursa through the failed abductor repair.
597. So far as the histopathology is concerned I regret to say that I am unable to rely on Professor Freemont's evidence in this and any other lead case, except to the extent that he agreed with Professor Nelson. It pains me to have to make such findings about any expert, particularly one who appeared to have an earnest desire to assist the Court, but Professor Freemont was an unreliable witness. He was evasive when challenged and produced convoluted explanations in an attempt to address the discrepancies between what he had written in his reports and the features of individual cases that were put to him in cross-examination. He appeared at times to be going to extraordinary lengths to support the claimants' case or his own opinions, even when it was demonstrated that he could not possibly be right. I found his evidence frequently confusing and often contradictory. I reached the view, ultimately, that he had persuaded himself that he was right and honestly believed himself to be right, and that he had become so convinced by this that he refused to countenance any possibility

that he might be wrong. He therefore lacked the necessary impartiality. Professor Nelson was his opposite in virtually every respect.

598. In all the lead claims, Professor Nelson had used a more powerful microscope to examine the slides; Professor Freemont explained that he did not have access to the same equipment. At one point, when Professor Nelson demonstrated that certain of the slides seen under more powerful magnification did not show what Professor Freemont claimed to be visible under lower magnification, Professor Freemont suggested that Professor Nelson had enlarged the wrong portions of the original images. That would have defeated the whole purpose of the exercise, and was patently incorrect.
599. In reaching his conclusions, Professor Freemont took into account features which were outside his expertise, or which were irrelevant, or articulated theses that were not supported by scientific studies, including the studies he relied on in support of them. For example, he sought to rely on the colour of the cytoplasm of macrophages, which he said was consistent with the presence of nanoparticles and a “hallmark” of ARMD. None of the experts, including Professor Athanasou, had previously suggested that the colour of the cytoplasm was even a non-specific feature consistent with ARMD in either the histological or the clinical sense of that term.
600. When Mr Antelme pointed out to Professor Freemont that his cytoplasm thesis was based upon what he had described as a “brown-grey” colour and that all the papers he relied upon refer to a different colour, namely blue or blue-grey, he sought to argue that if you looked at the papers you would see that the colours described as grey or blue-grey by the authors were actually brown. However, one would expect the authors of the papers to have given a more accurate description of the colour that they saw under the microscope than someone looking at reproductions of the slides in a copy of their work. Professor Freemont was eventually driven to accept that none of the papers suggested that the colour of the cytoplasm (irrespective of whether it was brown, grey or blue) was a hallmark or even an indicator that there had been an adverse reaction to particles of metal in the body. The most that they suggested was that the colour of the cytoplasm was due to nanoparticles of cobalt or chromium. There was nothing in them to suggest any link with ARMD.
601. There are so many deficiencies in Professor Freemont’s evidence that I cannot have any confidence in the conclusions that he reached. Much of his evidence emerged for the first time in cross-examination, despite the fact that he had produced a supplemental report and contributed to the joint report, and had even taken the opportunity to add annotations to some PowerPoint slides produced by Professor Nelson. In the course of his oral evidence, he produced another report overnight. This was all done in an attempt to justify various aspects of his report which had been demonstrated by Mr Antelme to be lacking in forensic rigour (or simply wrong).
602. It was common ground that in Mrs Blake’s case there were no signs of ALVAL (indeed, it was no part of her pleaded case that there were). I bear in mind the evidence of Professor Athanasou, to which I referred in paragraph 57 above, that on the rare occasion when ARMD has been found without a lymphocytic response, it was associated with heavy necrosis, macrophage infiltrate, and a large pseudotumour. Mrs Blake’s case had none of these features. I unhesitatingly accept the evidence of Professor Nelson that there is nothing observed on the slides of the tissue taken from



Mrs Blake that would support a finding of ARMD in the histopathological sense, let alone the support the clinical diagnosis of ARMD. His evidence is consistent with the content of the contemporaneous histopathology report. Although macrophages were present, which in and of itself is not unusual (as Mr Whitwell accepted), there is no evidence of an exuberant macrophage response and no significant inflammation.

603. There is no evidence of tissue necrosis. Whilst it was common ground that under extremely high magnification it was possible to see some dead macrophages, there were too few of them to suggest that they were being killed off by something (let alone by metal particles) rather than simply reaching the end of their natural lifespan. Most visible macrophages were viable. Professor Nelson provided detailed images which demonstrated the absence of areas of the synovial lining and the presence of a layer of fibrin but did not show any necrosis. Tissue identified by Professor Freemont as an area of cellular necrosis beneath the absent synovium was demonstrated by Professor Nelson to be viable fibrotic tissue. He was firm in his view that what he observed in terms of dead or dying cells fell short of what he would describe as tissue necrosis and put the case “very low on the spectrum”.
604. In the case of Mrs Blake, as in others, Professor Freemont was adamant that he could see extensive cell necrosis or tissue necrosis on the slides, drawing conclusions from certain features which he was at pains to point out to the Court using a cursor on the screen. When this evidence was put to Professor Nelson in cross-examination by Mr Preston, his courteous but emphatic response was that it was not possible for him to say on oath that the slides showed a quantity of cells that were dead or in the process of dying, for the simple reason that they did not show that. He pointed out that the images were taken in cross-sections or “layers” of tissue, which restricted the ability to deduce what was going on in the layers above and below. He added that the images relied upon by Professor Freemont were at far too low a magnification to draw any conclusion of this nature from them, which is why he went to a higher power of magnification to check. The effect of his evidence, though he was far too polite to say so in terms, was that no reasonable histopathologist could draw the conclusions that Professor Freemont had drawn from what was visible on the images that he was observing. Far from being necrotic tissue, it was viable tissue. He then gave a very clear and plausible explanation of why that was.
605. The items identified by Professor Freemont as large quantities of solid black metal debris were shown, when the images were magnified, to be light coloured, translucent and crystalline in appearance, which is inconsistent with their being metal particles. Professor Freemont was unable to explain the translucency if they were metal, and avoided addressing that topic when the point was raised in cross-examination. When subjected to polarised light, the particles were found by Professor Nelson to be refractile and birefringent, which again points away from their being either metal particles or products of corrosion.
606. Professor Freemont appeared to accept that the particles were refractile, and that generally cobalt and chrome are not birefringent under polarised light, although he said that some metal particles can be. However, he then speculated that the particles may have had an oxide coating and that this may have caused birefringence. The paper on which he relied for that suggestion provided no support for a finding that cobalt or chromium would appear birefringent when coated with an oxide and looked at under polarised light. The study in question related to structures produced by the

authors from silver, coated in iron oxide and subjected to a beam of light, and the reflected light was measured using instrumentation, rather than looked at under a microscope. By this stage in his cross-examination he was plainly clutching at any straw, however flimsy, to support the views to which he had become immutably wedded.

607. Professor Freemont also accepted that corrosion products are neither birefringent or refractile. The only way he was able to explain Professor Nelson's findings was to suggest that Professor Nelson had gone about assessing birefringence in the wrong way – a suggestion that was so absurd that Mr Preston wisely chose not to put it to Professor Nelson in cross-examination.
608. I conclude that Mrs Blake did not have ARMD. There is no evidence to suggest that she did. In those circumstances it is unnecessary to consider the effect upon her of revision surgery, because even if I were wrong in my conclusion that the product is not defective, Mrs Blake has not been able to establish causation, and her claim must fail.

#### **5.4 SYBIL STALKER**

609. Mrs Stalker was born on 25 June 1930. She underwent a left total hip arthroplasty on 2 May 2007. She subsequently underwent a right total hip arthroplasty using a Pinnacle MoM prosthesis on 17 November 2007. She was then aged 77. The latter prosthesis was revised on 15 December 2011 when Mrs Stalker was 81 years old. The liner and head were exchanged. The cup and stem remained. DePuy accepts that Mrs Stalker had a reaction to metal debris, but denies that it was an adverse reaction, or that she required a revision as a result of an ARMD.
610. Mrs Stalker, who is now aged 87, was too unwell to give evidence orally at trial, even via a video link. Her witness statement, signed on 6 December 2016, was therefore admitted under the Civil Evidence Act, but its contents were not agreed. The Court heard evidence from the orthopaedic surgeon who performed both the original and the revision operation on Mrs Stalker, Mr Antoni Nargol, to whom reference has been made earlier in this judgment. Mr Nargol is an experienced surgeon specialising in hip replacement and revision surgery. He has performed many MoM revision operations – he said, over 400 in the last 8 years. It came to light at a much later stage in the trial that he was the most prevalent of the “outlier” surgeons identified in the NJR statistics. He contributed to some of the scientific research papers that were referred to in this litigation.
611. Mr Nargol has been called as an expert on behalf of claimants suing DePuy in other jurisdictions, and he has also acted as a paid consultant to claimants in MoM litigation. These matters were properly disclosed in his witness statement. However, he failed to disclose, until it was elicited in cross-examination, that he had also acted as a paid consultant for claimants in the current litigation (though not for Mrs Stalker's case). I was not unduly concerned by this, even though he should have mentioned it, but unfortunately it was far from the only point of justifiable criticism that was made by DePuy about his evidence.
612. Mr Nargol's witness statement contained a transcription of a copy of his handwritten note of the revision operation, which had been overwritten by an unidentified person.

He said in his witness statement, signed in the usual way with a statement of truth, that he had “*inspected the original operation note*”. In juxtaposition with the transcription in the witness statement, this reinforced the misleading impression that the transcription was an accurate one. In later correspondence, Mrs Stalker’s solicitors said that Mr Nargol had been dissatisfied with the overwritten copy document and had retrieved the original handwritten note from the Nuffield Hospital before he signed his statement.

613. On the morning on which he was due to give evidence, a copy of the original operation note, minus the annotations, was produced. Mr Nargol explained that he had noticed that the transcription of the note in his witness statement did not tally with the original operation note. He brought this to the attention of Mrs Stalker’s solicitors. In his evidence in chief, Mr Nargol amended his statement to set out what he said was an accurate transcription of that note (which included words that had been omitted). If he had gone to the trouble of retrieving the original document to satisfy himself of what he had recorded, and had used it for that purpose at the time, it is difficult to understand how Mr Nargol came to allow an uncorrected, inaccurate and incomplete transcription of it to remain in his statement when he signed it.
614. It is to his credit that Mr Nargol drew the discrepancies to the attention of Mrs Stalker’s solicitors, which led to the original operation note being produced, and the statement being corrected. However, it is troubling that the only explanation he could give for the fact that, despite his checking the original, the inaccurate transcription was included in a witness statement to which a statement of truth was attached, was that “*the writing that I’ve written in the operation note is really poor and it is difficult to read.*” That was manifestly incorrect: unlike the note of the primary operation, the handwriting is relatively easy to decipher, (even by someone other than the author), and is certainly more legible than the copy with the overwriting on it.
615. In any event, it does not afford an explanation, as Mr Nargol was plainly able to decipher the original sufficiently to appreciate the inaccuracy of the transcription in the witness statement shortly before he came to give his oral evidence. Either he did not check the original operation note before he made his statement, as he said he had, or any check was so perfunctory that he did not appreciate that the transcription was inaccurate, or he did not consider the differences to be worthy of correction. The transcription in his witness statement of his note of the primary operation was also inaccurate, though the error did not materially affect the gist of what was recorded. At its lowest, Mr Nargol was careless about the content of a witness statement to which he was prepared to append a signed statement of truth. That led me to have concerns about his appreciation of the importance of accuracy, both in terms of making records and in terms of giving evidence.
616. The inaccuracy of the original transcription and the attempt to explain away the inaccurate witness statement pale into insignificance, however, when compared with Mr Nargol’s general presentation as a witness, which was markedly different from that of the other operating surgeons who gave factual evidence. When faced with contemporaneous evidence that did not support Mrs Stalker’s case, Mr Nargol was evasive or tried to explain it away; he also sought to bolster evidence he plainly believed to be helpful to her. In either case, he sometimes volunteered explanations that flew in the face of the contemporaneous documents, and that were not in his witness statement.

617. Mr Nargol's witness statement is dated 8 December 2016. In it he stated that he had some independent recall of Mrs Stalker's case but in the main he was reliant on what appeared in the medical records. He said that where he had a specific recollection he would make it clear. Unfortunately, on those occasions when he sought to draw on such "specific recollections" in his oral evidence, on the most benevolent interpretation of that evidence, he appeared to be reconstructing what he thought must have happened. He was dismissive of any views that did not coincide with his. He was also prone to expressing himself in overly dramatic terms.
618. It may be that Mr Nargol's behaviour in the witness box was, at least in part, driven by a desire to defend his decision to operate on Mrs Stalker, and his diagnosis of ARMD on the flimsiest of grounds, but that does not excuse it. I found Mr Nargol to be, in many respects, a witness on whose evidence the Court could not safely place any weight. His unreliability has complicated the Court's task of ascertaining whether, on the balance of probabilities, Mrs Stalker did suffer from ARMD. However bad an impression he made in the witness box, it does not mean that he is not a good surgeon.
619. Matters were not made any easier by the fact that the claimants' orthopaedic expert in this case was Professor Kay, who, whilst undoubtedly very knowledgeable, and generally quite fair, on a few occasions descended into the arena as an enthusiastic advocate for his clients' case. At times, he tried to anticipate the questions in cross-examination and get his retaliation in first. For those reasons, I was driven to treat his evidence with some caution, although most of the time he made concessions where appropriate and gave sound reasons to justify his opinions. I had no similar reservations about any of the other orthopaedic experts, irrespective of who called them, though Mr Smith's oral evidence was an improvement on his written reports. Of all the lead claims, I found this by far the most difficult to determine.
620. In May 1995, Mrs Stalker had a fall and suffered a sustained fracture to the neck of her right femur. This was initially fixed with cannulated hip screws and subsequent X-rays showed that the fracture had united and the screws were in a good position. She was told that there was no need to remove the screws unless they caused trouble.
621. X-rays taken of Mrs Stalker's pelvis in late 2006 following complaints of severe hip and groin pain, particularly on the left side, indicated that she had a degree of bilateral osteoarthritis of her hips. Her GP noted that she was now at the stage where this was threatening her general independence and mobility. She would consider an operation if offered, as painkillers were having no appreciable effect.
622. On 16 February 2007 Mrs Stalker's GP wrote to Mr McMurtry, a consultant orthopaedic surgeon at the James Cook University Hospital, explaining that although Mrs Stalker had been placed on another surgeon's list to have her left hip replaced on the NHS, she was in such severe pain that she was not prepared to wait, and had asked for a private consultation. Mr McMurtry saw Mrs Stalker on 15 March 2007. He noted that over the past six months she had had increasing pain from the left hip which was starting to significantly impact on her lifestyle. She did not go out very much at all and was severely limited in her walking distance. She had an obvious limp on the left. Her right leg was about 1½ cm short, secondary to the hip fracture. Mr McMurtry went through the pros and cons of an arthroplasty with her, and she was quite keen to proceed as soon as possible.

623. The operation was carried out on 2 May 2007 using a standard Exeter cemented hip prosthesis. Mr McMurtry's letter to her GP following the seven-week review indicated that all was well. He said that the wound had healed nicely and that her arthritic pain had resolved. Mrs Stalker had been fitted by a physiotherapist with a heel raise on the right shoe, and she was now mobile with one stick with reasonable balance. She was going up and down stairs unaided. Mr McMurtry noted that the right hip was still quite painful, with no internal rotation. He told the GP that he was sure that Mrs Stalker would find that her right hip pain was continuing to limit her functionality, and he had advised her to think carefully about whether or not she wanted to proceed with having the right side done.
624. Three months after the left hip arthroplasty Mr McMurtry reviewed Mrs Stalker and said that she had rehabilitated quite nicely, but was slow because of her arthritic right side. She had asked Mr McMurtry about removal of the screws to try and settle the pain on the right side, but he had advised her that he did not feel that this would really help. He had advised that the best option to settle her symptoms would be to have another hip replacement, which would also address her leg length discrepancy. However, having just come through one hip operation Mrs Stalker was not too keen to proceed with another.
625. Matters were left as they were until 18 September 2007, when Mrs Stalker was seen in the orthopaedic clinic of the Cleveland hospital by Mr Nargol. She had severe pain in her right groin and thigh. Her right hip was very stiff. The X-ray showed that the cannulated screws were in place but there was a collapse of the lateral part of the femoral head with osteoarthritis. Mr Nargol said that he did not believe that just simply removing the cannulated screws would work, and that she needed a total hip replacement using a Corail stem. Mr Nargol explained that his plan would have been to use a MoM configuration because Mrs Stalker had a smaller anatomy and therefore needed a 50mm cup. He agreed that she would not have been able to use a 36mm polyethylene liner at that stage. In re-examination, he said that he could have put in a 32mm MoP articulation. Of course, that would have meant using a smaller sized femoral head, increasing the risk of dislocation (which, it will be seen, was something that Mrs Stalker was really worried about).
626. Unfortunately, the consent form signed for the right hip total arthroplasty on 17 November 2007 is virtually illegible. There is no evidence as to what she was told about the choice of hip. Mr Nargol's operation note is also difficult to read. However, it records that he used a 50 mm Pinnacle cup with a 36mm metal liner. The screws were difficult to remove because a Stryker screwdriver had been sent in error. There was a bone graft to the screw holes.
627. Mrs Stalker's recovery took some time, but there were no symptoms giving rise to any cause for concern. Around six weeks after the surgery, in January 2008, she was described as having "reasonable" abductor muscle control. That was something that Mr Nargol, who had used the Hardinge approach, said he would have checked for, particularly to ensure that the repair to the abductor muscle had not failed. He said that he would have expected some weakness of the muscle at that stage. At her follow-up consultation three months after the operation (on 20 February 2008) she was described as having no pain, no stick and good muscle power. Mr Nargol explained, and I accept, that this meant that she was not using a stick and did not have a visible limp.

628. On 19 August 2008 Mr Nargol wrote to Mrs Stalker's GP stating that she had done well with both her hips. She had no discomfort or problems at all, and her wounds had healed beautifully. The plan was to see her only if she had any problems in the future. An assessment of Mrs Stalker by Social Services on 3 December 2008 indicated that she could get around indoors and could manage stairs; she could do light housework but needed help with heavy work, and she needed someone to go with her to help her when she went shopping. It appeared that the operation had achieved what it had set out to achieve in terms of restoring Mrs Stalker's mobility and freeing her from pain.
629. There is nothing in the medical records to indicate that Mrs Stalker made any complaint of further pain with her hips on either side before she received a standard letter from Nuffield Health dated 24 June 2010 that was sent to all patients who had received a MoM hip joint replacement, in the wake of the MHRA guidance. The letter said that it appeared that a small number of patients with MoM articulated hip joints had developed problems after their surgery, potentially due to metallic wear, which in some cases was not consistent with the outcomes expected following hip joint replacement. These problems could present with pain and reduced mobility. The number of patients experiencing problems was less than one in a hundred, and therefore in most cases there would be no issue. The letter explained that all patients who had had such an implant should receive an annual follow-up appointment for the first five years following the implant. Mr Nargol's evidence was that the letters were sent to different cohorts of patients, beginning with those who had received ASR MoM hips, then patients who had received Pinnacle MoM hips, and finally other patients who had received different MoM hip replacements.
630. Mrs Stalker initially ignored the letter, which arrived on her 80<sup>th</sup> birthday. However, she was contacted again in early 2011 as part of the follow-up process, and on this occasion she kept the appointment. Mrs Stalker was seen in Mr Nargol's clinic by a nurse practitioner on 18 January 2011. She recorded that Mrs Stalker's Harris hip score was 91 (out of a possible 100), which was classed as "excellent" under that system. Her UCLA score was also good. She had no hip pain and scored the maximum on her range of motion, and although she was recorded as only being able to walk 200-300 metres, that was consistent with her presentation prior to the surgery. All these findings were very satisfactory, and Mrs Stalker was very satisfied with her hip. If the MHRA guidelines had been followed, Mrs Stalker required no further investigation, since her prosthesis had not been causing her any pain and it was functioning really well.
631. Despite this, Mrs Stalker's blood ion levels were measured, and the results showed slightly elevated levels of chromium (7.5 ppb), marginally above the MHRA screening threshold, although the cobalt level was 5.9 ppb and thus below that threshold. Mr Nargol reviewed the test results (which he described at the time as "surprising") and the results of X-rays, which he described as excellent, and instead of ordering another blood test three months later, he decided to order an ultrasound of the right hip immediately, to check for fluid. In his oral evidence, Mr Nargol explained that this approach was taken in the Nuffield Health Tees centre for all patients with Pinnacle MoM implants even if they were wholly asymptomatic, if the cobalt levels were higher than 4 ppb. That centre adopted a more cautious approach than the MHRA screening guidance. His personal view was that normal blood metal

ion levels in a well-functioning hip prosthesis were 2ppb, and the MHRA guidance was wrong.

632. So, despite the fact that Mrs Stalker was experiencing no pain and no other problems with her right hip, and professed herself happy with it, it was decided to carry out further investigations. Mr Nargol told the Court that this was purely based on the fact that she had a Pinnacle MoM hip. He also said that he made a note that her cup size was 50mm because his own personal experience by then was that the 50mm cup sizes seemed to be failing at a higher rate than larger prostheses. Yet the MHRA guidance at the time stated that the size of cup was relevant only to resurfacing. Mr Kim, DePuy's orthopaedic expert in this case, observed that a link with cup size was not established by evidence in the orthopaedic literature, and Ms Smith found that the size of the cup was not statistically significant once gender and age are taken into account.

633. The results of the ultrasound, carried out by a radiologist named Dr Raju, reported that there was:

*“abundant echogenic joint effusion noted in the lateral and posterior lateral aspect of the right hip (above and posterior to the greater trochanter) but there was hardly any fluid collection (echogenic or anechoic) noted in the anterior or anterior medial aspect of the right hip. The right sided ileopsoas tendon had normal appearances. The tendons of the gluteus medius and minimus had decreased reflectivity but were intact on the lateral aspect of the right hip”.*

The ultrasound report is extremely confusing because the finding of “abundant echogenic joint effusion” is at odds with the description of where the effusion is found. As Dr Wilson pointed out, the joint does not extend to the trochanteric region. It is also inconsistent with the finding that there is hardly any fluid collection in the anterior or antero-medial aspect of the right hip.

634. Dr Ostlere and Dr Wilson agreed that the appearances on the (still) ultrasound images that they were able to review were in keeping with an extra-articular collection (fluid within a cavity outside the joint) on the lateral side of the greater trochanter extending to the proximal femur, but that a communication between the collection and the hip joint is not visible on the imaging. Far from being “abundant”, there was no detectable joint effusion on the ultrasound. Dr Wilson's view was that although the images were in keeping with an extra-articular collection, they could also be interpreted as showing an area of boggy tissue. It is possible that Dr Raju, having seen a collection in that area, assumed that there must be a connection with the joint. If he subscribed to Dr Ostlere's view that a collection round the trochanter in a patient with a MoM prosthesis means ARMD unless and until proved otherwise, that is a highly likely scenario.

635. The expert radiologists also agreed that the images poorly demonstrated the tendons but drew different conclusions from what they could discern. Dr Ostlere's view was that the state of the gluteal tendons could not be assessed on the scan because of the orientation of the tendon (oblique to the probe) which would result in artificial reduction in reflexivity. Dr Wilson's view, which I prefer, having regard to his explanation for them by reference to the images, and the view expressed by Dr Raju about them at the time, was that the images showed decreased reflexivity in both the

gluteus medius tendon and part of the gluteus minimus tendon, consistent with chronic tendinopathy.

636. Mr Nargol did not query the apparent inconsistencies in Dr Raju's report or view the images himself; he latched on to the word "abundant". He said in his evidence that this kind of language scared him and that a revision would be likely in the light of it. Dr Raju was described by Mr Nargol as "*one of the world's top radiologists for scanning hips like this*". He said that from his knowledge of Dr Raju's terminology "abundant" was his highest level and meant "*an enormous amount*". To an extent, therefore, Mr Nargol may have been led astray by his radiologist, whose confusing statements he made no attempt to clarify, even though it appeared from his evidence that the two of them had regular discussions and worked closely together. This probably conditioned his pre-existing expectations of finding Mrs Stalker to be suffering from ARMD. If he had sought clarification, he should have discovered that Dr Raju had found no fluid in the joint itself, and that no communication had been identified, but I very much doubt if that would have made any difference to his actions.
637. On 23 May 2011 Mr Nargol wrote to the GP saying that "*this lady has no problems with her hip whatsoever, but I am concerned that her blood tests showed a cobalt of 5.9 and chromium of 7.5. Her ultrasound scan of her right hip shows abundant fluid. I think this lady is developing an ALVAL type problem. I plan to aspirate her hip and I have repeated her chromium and cobalt levels today.*" At that stage, Mr Nargol was already thinking about revision surgery.
638. Mrs Stalker's witness statement says that from around early summer, probably June 2011, her right hip started to get stiff and painful, she had great difficulty walking because she felt she was dragging her right leg, and she had to stop and rest after walking around 100 yards. She found bending extremely difficult and painful. She also stated that the pain made it difficult for her to sit and concentrate on driving. There is no mention of any of these symptoms in any contemporaneous medical note at any time prior to the revision surgery in December 2011.
639. When she was examined by Professor Kay, Mrs Stalker gave him a substantially modified account. She said that "*by mid-2011 she had started to develop some discomfort and stiffness*" and that the pain in her right hip had increased significantly by the time of the revision (in December). Her walking distance had decreased, and she was now having to use a walking stick to get about, as well as needing help at home. Professor Kay said that Mrs Stalker did not appear to be confused. He plainly took what she was telling him at face value. However, the medical records do not tally with either version of her symptoms in this period. Mrs Stalker has probably persuaded herself, after the passage of many years, that she was feeling pain and experiencing these difficulties at that time, but I find that her recollection is mistaken.
640. Mr Nargol's manuscript note of the aspiration, which was carried out on 5 September 2011, records that 15ml of "*yellow but thin*" fluid was extracted. The next passage was transcribed by Mr Nargol as reading "*sent for cobalt and chromium and culture and sensitivity*". By this he meant that he was going to have the fluid around the hip tested for metal ion levels. It ended with a note that he would see the patient in six weeks for revision *if painful*. He would not have made that note if the patient had mentioned to him that she was already in considerable pain. If Mrs Stalker was in



much pain at the time as she told Professor Kay she was, she would not have remained quiet about it when Mr Nargol carried out the aspiration.

641. Mr Nargol's oral evidence was that he arranged the metal ion tests of the aspirate to see if they were "in sync" with the levels of metal ions in the blood. He accepted that there were not at the time, and there are not now, any systems or criteria to consider ion levels in hip fluid as a threshold for management of a patient with a MoM hip articulation. Mr Herlekar, who operated on Mr Woods, agreed that it would be unsurprising to find metal ions in hip fluid with a patient who had a MoM hip prosthesis, and that there are no levels or guidance as to what to expect in hip fluid in terms of ions in such a patient.
642. Mr Whitwell said that this was not an investigation that he was familiar with or would undertake himself; he was aware that some papers had talked about it, but he had not heard of any guidelines by which to assess the extent to which ion levels in aspirate might correlate with any reaction. Professor Pandit said in his expert report that although isolated reports do exist in literature about an association between metal ion levels around a MoM hip and ARMD, no long-term large studies have been carried out to establish thresholds for revision of MoM hips for suspected ARMD based on the metal ion levels in intra-articular fluid. There are no recommendations that suggest that measuring such levels is an appropriate method of monitoring or assessing a patient. His opinion, which was not challenged in cross-examination, was that "*there is no reliable data to assist clinicians on interpreting metal ion levels in intra-articular fluid*". So, although the level of metal ions in the aspirate were found to be higher than the level measured in Mrs Stalker's blood, that tells one nothing about the presence or absence of ARMD.
643. Mr Nargol told the Court that 15 ml was not a normal amount of fluid to extract. He said that if you aspirate 20mls of fluid it is highly correlated with soft tissue damage and that 15ml was "*getting towards that*." However, he then rapidly qualified that evidence by volunteering that there were "*not a lot*" of publications linking the volume of fluid with soft tissue damage (in fact, despite the numerous scientific papers adduced in evidence in this trial, no-one cited any linking any particular volume of aspirate with soft tissue damage). Mr Whitwell said he would not expect to extract much aspirate from a well-functioning hip. He agreed that there would be a range of volumes on aspiration, and that 10ml would be "*quite normal*". His evidence as to whether 15ml would be a cause for concern was confused, as he initially agreed that it would not be, and then on reflection said that it might. Professor Kay said in his oral evidence that if the hip was completely well fixed and there was no evidence of loosening on the scan, which was the case here, 15/20ml is a reasonable amount.
644. Of course, this presupposes that the fluid was extracted from the hip joint itself, and not from a collection around the trochanter. One of the oddities in this case is that there is a single, high grade radiographic image which shows the needle stopping well short of the joint. Professor Kay described how a surgeon would see a series of low-grade images as he was inserting the needle. He agreed that the normal practice would be to ask the radiographer to record the image when the needle was in its final position. Dr Ostlere said he assumed that the image in Mrs Stalker's case was taken simply in order for the surgeon to make sure he was correctly aligning the needle. Unfortunately, Mr Nargol was not asked about that image and so he had no chance to explain it, but Dr Ostlere's assumption makes sense in the context of Mr Nargol's

purpose in carrying out this procedure. There would be no point in an orthopaedic surgeon who is seeking to aspirate a hip joint deliberately leaving the needle short of the joint and drawing the fluid from somewhere else. That was Professor Kay's view, and I agree with it. I am not prepared to find that a surgeon of Mr Nargol's experience would have done such a thing if he was trying to aspirate the hip joint. He would know if he had made contact with the joint or not.

645. Mr Nargol said that yellow fluid is "*usually a sign of ALVAL and it is a poor thing.*" This was a view he expressed in his letter to Mrs Stalker's GP of 8 November 2011. He repeated this view several times in his evidence; indeed, at one point he said that many centres regarded yellow fluid as "*synonymous with ALVAL*". None of the expert orthopaedic surgeons agreed with this, and there was no other evidence which would support it. Professor Pandit said in his expert report that he did not understand this and had not encountered any literature which would support such a conclusion. Surgeons will often note the colour of fluid at revision and reports of green or brown fluid have been made; the findings may be variable. The normal colour of synovial fluid is yellow (straw-coloured) and therefore it might have been thought that the presence of yellow fluid was either neutral or reassuring. Mr Whitwell agreed; he had seen a whole spectrum of colours in aspirate, which he presumed would be coming from the haemoglobin or metal staining.
646. Professor Kay made no comment about these matters in his expert report; in cross-examination he said he did not think the colour of the fluid mattered. Mr Kim said in his expert report that so far as he was aware there was no evidence in the orthopaedic literature to support the assertion that yellow fluid was a poor prognostic indicator. However, all the expert reports were written before Mr Nargol interpreted his note in the witness box. In cross-examination, he said that by "*yellow*" he actually meant "*fluid you can't see through*". It is odd, to say the least, to use a colour to describe opacity, and I got the distinct impression that when Mr Nargol gave that evidence, he was casting around for an explanation to overcome the problem that no expert in the case agreed with his view about the significance of the colour. Mr Whitwell, who gave evidence after Mr Nargol, was asked about this topic in cross-examination, and was firm in his view that even if "*yellow*" meant what Mr Nargol said it did, yellow is not a poor prognostic indicator. He did not suggest that any inference could be drawn from opacity.
647. Professor Kay told the Court that when he first became involved in revising metal on metal hips, there would often be a milky, cloudy yellow fluid which looked like dilute pus and it would be sent off to the lab to look for signs of an infection. Whilst I have no reason to doubt that evidence, the description of such fluid does not easily correlate with the phrase "*yellow but thin*" used in the note of the aspiration. Even if that was what Mr Nargol was describing, Professor Kay did not go so far as to suggest that any clinical significance could be attached to its presence.
648. I accept Mr Whitwell's, Mr Kim's and Professor Pandit's evidence in preference to that of Mr Nargol, who appeared to me to be going to some lengths to exaggerate the importance of the alleged opacity and amount of the fluid he aspirated, which I find were and are of no significance in determining whether this was a case of ARMD. I am not prepared to accept that when he wrote "*yellow but thin*" he meant anything other than what those words would normally convey to any reader.

649. Mr Nargol saw Mrs Stalker again on 21 September 2011. The latest blood ion tests showed the levels were elevated from the previous tests – Cr was now 10.7 ppb and Co 9.1. Since the cup and head appeared well positioned on the X-rays, he surmised that there might be a problem at the taper end, and that is what he suggested to Mrs Stalker. His notes recorded that he had a “long chat” with Mrs Stalker and her son. There is nothing in Mr Nargol’s notes to suggest that during that conversation, Mrs Stalker told him that she was in any pain, or that she was dragging her leg or had difficulty in walking. He would have been able to see if Mrs Stalker was limping or finding it difficult to walk or to sit for any length of time in his consulting room. Given that he was specifically thinking about revision if the hip was painful, I find it inconceivable that Mr Nargol would have failed to ask his patient on this occasion whether she was in pain, and noted the response if she was. A deterioration in Mrs Stalker’s function or increased pain would have been critical information to have considered. A consultant, however busy, would normally record his patient’s symptoms, and he did make a note of the consultation.
650. Mr Nargol ordered further blood ion tests which showed that the ion levels had come down again to 8.7 ppb (Co) and 7.1 ppb (Cr). By the time of the next consultation, on 7 November 2011, Mrs Stalker’s disabled son, for whom she was the carer, had passed away. Mr Nargol said that (in the light of this) she need not rush into surgery immediately, but Mrs Stalker told Mr Nargol that she did not want to wait any longer and that she wanted to have the hip sorted out as soon as possible.
651. Mrs Stalker told Mr Kim when he examined her that she had read about MoM hip replacements extensively and was very worried about metal particles disintegrating in her body. She said that she had read a lot of articles in the press regarding MoM hips and she was very unhappy with the fact that she had a very satisfactory left hip replacement and was left with a right hip replacement which was failing. This concern was bound to have had some influence on the decision to revise, though Mr Nargol did not appear to need encouragement. As Mr Kim said, *“if a patient is very concerned, the reality is that for a surgeon, you don’t want to go against what your patient feels”*. In this case, Mrs Stalker’s views coincided with Mr Nargol’s.
652. On the basis of all the pre-operative information, Mrs Stalker was revised because Mr Nargol applied a very low threshold for investigation and revision, and because he relied on symptoms that were not indicative of ARMD. The only feature that was potentially consistent with such a diagnosis was the collection of fluid around the greater trochanter, but ARMD was not the only explanation for that. Mr Nargol did not follow the MHRA guidelines. He agreed that if it were not for the fact that he took a special approach for patients with Pinnacle MoM prostheses, Mrs Stalker would have had no blood tests and no investigations. A surgeon following the MHRA guidelines would not have investigated, let alone revised Mrs Stalker.
653. Mr Oppenheim submitted that even if the actual reasons for revision were unjustified, it did not matter because when Mr Nargol performed the surgery, it was apparent that his suspicions that she had ARMD were well-founded. That depends on whether what was found during and after the operation supports the diagnosis.
654. The revision surgery took place on 15 December 2011. Mr Nargol replaced the metal head with a ceramic head, and the liner with a polyethylene liner. The original operation note reads as follows:

*“pre-op HHS (Harris hip score) 91      UCLA 5  
chromium 8.7 cobalt 7.1  
revision of right, Corail, 50 Pinnacle two ceramic head +9 millimetres for ALVAL  
severe ALVAL, abductors 70% gone, lifted off  
very large pseudotumour 15 cm x 10 cm. Awful.  
external rotators weak  
as much of pseudotumour excised, liner removed with now a 50 polyliner +4  
stable, ceramic to pre-revision head +9  
short pre-op.”*

655. It is significant in this context that in his operation note, Mr Nargol referred to the high Harris hip score and UCLA score taken in January 2011. He sought to explain this on the basis that he had been too busy to do another set of tests. However busy Mr Nargol was, he would not have put scores which were by then almost a year old in his note, if he had ascertained that his patient’s mobility had significantly deteriorated in the intervening period and that she was currently suffering from debilitating pain. He accepted this when it was put to him by me.

656. Mr Nargol sent off large samples of the excised tissue for histological analysis. The report of the histopathologists who carried out the examination at the time stated as follows:

*“Sections from the peri-prosthetic tissue show surface covered by fibrin and no intact synovial lining is present. Some of the tissue shows villiform projection. The underlying tissue shows marked hyalinisation. Hyalinisation of the vessel walls is noted. There are no active germinal centres, or distinct granulomas. There is moderate infiltrate of lymphocytes mixed with histocytes present as scattered cells and also in perivascular distribution. Pigment laden macrophages are scattered throughout the specimen. The metal particles are scanty and barely visible (particle load 1+).*

*The appearances are consistent with moderate chronic ALVAL.*

*Tissue from right hip: consistent with moderate chronic inflammation – ALVAL”*

It is accepted by DePuy that there are some findings that *could* be consistent with ARMD. However, the key feature absent from the contemporaneous histological reports is tissue necrosis.

657. It was agreed by the histopathology experts that the first sentence of the report contains a mistake; some areas of the synovial lining were intact. In any event the absence of synovial lining is not a histological feature of ARMD, though it is consistent with, among other things, an ALVAL reaction. The presence of fibrin is also a non-specific finding. Professor Nelson’s evidence was that the immunological reaction seen and described by him as allowing for a score of “low to moderate ALVAL” on the Campbell scale (which is not a diagnostic) could be a non-specific reaction in an already abnormal osteoarthritic joint, even after the patient has undergone a hip replacement. It could also be a mild inflammatory reaction to metal particles, but that reaction would not necessarily have adverse effects on the patient.

658. That evidence was consistent with Professor Athanasou's evidence. He stated that synovial inflammation, fibrosis and villous hypertrophy can occur in osteoarthritis and that perivascular lymphocytes and lymphoid aggregates as well as a macrophage infiltrate can be seen in this condition, though the macrophage infiltrate is not as marked as in ARMD. He went on to explain that tissue necrosis is not seen in osteoarthritis, and that in ARMD the necrosis is "*often extensive and associated with metal wear particle deposition and a heavy foreign body macrophage response*". The histopathology in Mrs Stalker's case showed no heavy foreign body macrophage response, but a scattering of macrophages throughout the tissue and very little visible metal debris.
659. Professor Nelson's evidence, which I accept, is that there is no evidence of tissue necrosis. Professor Freemont's evidence that there was "*extensive necrosis*" (which the histopathologist at the time failed to note) was completely unreliable and I have no hesitation in rejecting it. That evidence was initially based on two relatively low magnification images which he sought, unconvincingly, to equate to the images in Mr Haley's and Mrs Emery's cases. He then sought to identify areas of "necrosis" on Professor Nelson's higher-powered images, but this was done by seeking to persuade the Court that it was possible to see individual dead or dying cells in isolation (I prefer Professor Athanasou's evidence that this is virtually impossible). Professor Freemont accepted that, if dead cells are visible under the microscope, it is impossible to distinguish between cells that had died naturally and cells that had been killed.
660. One of the images taken by Professor Nelson shows a layer of surface fibrin, underneath one area of which is the tissue that Professor Freemont described as having white spaces and comparatively fewer cells than other areas of the image. Professor Nelson pointed out that as this was a cross-section of tissue, some of the nuclei would be visible, but others would not, because the "slice" has not gone through the centre of each cell. In any event, tissue is made up of a variable number of cells. Thus, even if an area of tissue has fewer cells in it, that does not mean that the tissue is necrotic; in this slide, the histological features of necrotic tissue, namely, dead cells, and "ghost" forms of collagen fibres and blood vessels, are absent.
661. Professor Freemont's analysis of what was shown in Professor Nelson's slides was provided for the first time in his oral evidence, despite the fact that he had sight of the images before he prepared his supplemental report. He could also have commented on them in the joint statement. The fact that he failed to do so is telling. This, I find, was yet another example of Professor Freemont searching around for a means of rehabilitating his views when the evidence on which he originally relied was shown not to support them. Even if Professor Freemont had not been the maverick witness that he was, I would still have preferred Professor Nelson's explanation of these images, since he was the person who took them and was best placed to explain what he saw under the microscope.
662. Those macrophages that were visible on the enlarged slides were viable, and Professor Freemont was constrained to accept this when it was put to him. He then suggested that it was possible to see that the macrophages were beginning to die or as he put it "*that's a reflection of toxicity that's sublethal. So this cell may be dying but at the moment it's only badly injured*". Bearing in mind the evidence of the immunology experts (one of whom was the only toxicologist called to give evidence at trial) about the state of scientific research at present, this concept of sub-lethal

toxicity appears to be completely novel. What is visible on the slides appears to represent the normal response of macrophages to the presence of a very tiny amount of metal debris. This was a further example of Professor Freemont making up a thesis to try and salvage some credibility after being proved wrong. The point was not put to Professor Nelson in cross-examination, and for good reason.

663. Mrs Stalker needed a brace for six weeks after surgery. She made a good recovery. Her physiotherapy records on 12 January 2012 showed that she was increasing her mobility and function, though she was concerned about the risk of dislocation when she removed her brace. Those concerns were reiterated on 7 February, though she was reported to be progressing well and managing all functional tasks with the brace in position. The brace was removed by physiotherapists on 12 February 2012 and they assessed Mrs Stalker as progressing well. She managed exercises and mobility with no physical problems, although she was reluctant to move because of a lack of confidence. By her next physiotherapy appointment, she was recorded as mobile unaided with no disturbance, and as having no problems with mobility or function at home. On 29 March 2012 she was noted to be returning to driving. The physiotherapists noted that she had achieved a very good improvement and outcome and was managing all tasks and exercises with confidence, stability and strength. They therefore discharged her.
664. Mr Nargol reviewed Mrs Stalker on 9 May 2012 and observed that she was “*coming on slowly and steadily*”. Her blood metal ion levels were tested and subsequently reported to be very low. He next saw her on 2 February 2013. His letter to her GP refers to “slight pain” and a Harris hip score of 84/100. On 6 September 2013, the date of her next review, her Harris hip score was 85, which Mr Nargol accepted was classed as a good score. He said: “*considering the problems she has, it’s a very good score*”. In fact, it was not that much worse than the score of 91 recorded in January 2011. She was experiencing intermittent pain.
665. There is nothing of further note in Mrs Stalker’s medical history until 2017. In January that year, her solicitor asked her GP if she could be referred for physiotherapy to her hip, as she was becoming progressively stiff and less mobile. The GP examined her and recorded that she was still independently mobile and had only a minor limp. In a letter dated 18 January 2017 the GP stated that “*She is actually quite fit, well, active and mobile despite her age.*” A physiotherapist who saw her in February 2017 noted weakness in her quads, hip abductors and external rotators. She was given some exercises to try.
666. However, in May 2017 the physiotherapist noted that Mrs Stalker had not managed to do many of the exercises as she had been busy in her garden. In June 2017 she saw her GP who noted that she was independently mobile, with no stick. He was of the view that difficulties in turning fully to the right that she was complaining of to him were nothing to do with her hips, but more likely to be due to stiffness in her neck and shoulders.
667. On 3 October 2017 Mrs Stalker saw Mr Nargol, who reported that she had “*done ok until 5 weeks ago when she has had 3 painful movements of her hip*” when she was bending in her garden, which resulted in severe pain around her thigh, going up her body. He suspected that her hip was subluxing. He found that she had restricted movements, although she was mildly Trendelenburg positive on the right and left,

which he found surprising. She walked without support and with a moderate limp. He ordered an ultrasound scan.

668. Mrs Stalker underwent an MRI scan and an ultrasound. The images confirm that there has been damage to the abductors, as noted by Mr Nargol at the time of the revision surgery, but all the experts agreed that the images do not enable the cause of that damage to be identified. Dr Ostlere agreed that one could not tell from the imaging whether it was caused by ARMD or by a failed repair following the Hardinge approach.
669. The abductors were the only muscles that were noted to be damaged in the operation note. ARMD would not be discriminating in its choice of muscles to damage, as is demonstrated by the nature and extent of the damage experienced by Mr Haley. In contradistinction to Mr Nargol's evidence, both Professor Kay and Mr Kim described other soft tissues and muscles that could have been damaged by ARMD, such as the vastus lateralis, the extensors, the hip flexors and the external rotators. When it was suggested to Mr Kim in cross-examination, consistently with Mr Nargol's oral evidence, that when Mr Nargol wrote "*lifted off*" he was describing a pseudotumour bursting through and destroying the muscles, Mr Kim observed, with justification, that the operation record was poor. He explained that as a practising orthopaedic surgeon himself, when he says the abductor muscles are damaged it means "*it's off, it has come off, it has lifted off, it has gone*" and that was how he interpreted the note. In those circumstances it is impossible to draw any clear conclusions from the phrase "*lifted off*". I am not prepared to accept as reliable Mr Nargol's attempt in the witness box to reconstruct what he meant.
670. As Mr Kim observed, the damage appeared to be localised to the area that one would see if there was a failed abductor repair following a Hardinge approach. He could not reconcile the damage recorded on the face of the operation note with the description in it of "*severe ALVAL*" by a surgeon who would have "*seen it all*". Professor Kay accepted that the damage discovered at revision was entirely consistent with the damage being caused at the time of the primary surgery.
671. It is also of significance that Mrs Stalker continued to have good function and symptoms after revision, as demonstrated by her high Harris hip score, despite the extensive nature of the abductor damage noted by Mr Nargol and confirmed by the more recent images. Her presentation before and after revision surgery was very similar. There was a consensus between Professor Kay and Mr Kim that, whilst a patient who had a failure of an abductor repair following surgery using a Hardinge approach would be expected to exhibit signs of weakness following primary surgery, and to walk with a limp, there are some patients who are able to compensate functionally for failed abductors. Mrs Stalker obviously fell into the latter category.
672. I appreciate that a patient with a large pseudotumour may experience only low-level pain. Therefore, the absence of reported pain, as such, is a neutral factor. On the other hand, if the abductor damage was a result of ARMD progressively destroying Mrs Stalker's muscles in the period leading up to revision, both Professor Kay and Mr Kim would have expected to see a corresponding reduction in function and increasing pain in that period. Even if I had accepted Mrs Stalker's account of her symptoms in 2011 as given to Professor Kay, (which I did not, as it was contrary to the contemporaneous medical records) it would not have fitted that picture.

673. In short, the evidence in Mrs Stalker’s case points away from her suffering from ARMD, but for one aspect of the contemporaneous operation note. A very experienced orthopaedic surgeon, who had no reason not to record what he saw at the time of surgery, has noted that there was a “*very large pseudotumour*”. However, there are a number of reasons why I have come to the conclusion that that finding must have been mistaken.
674. First, Mr Nargol excised a large quantity of tissue and sent it off for histological examination, and although there were indicators of inflammation, there was no appreciable tissue necrosis visible within it. Professor Athansou’s evidence was that a pseudotumour was characterised by extensive cell and tissue necrosis. He accepted that there could be ARMD (in the histological sense) with minimal necrosis, (the parameters of “minimal” were never defined), but no expert suggested that there could be ARMD without any necrosis, especially if there was a pseudotumour. Soft tissue damage is an essential part of the clinical diagnosis of ARMD.
675. Moreover, the ALVAL type phenomenon observed was not the “high” level that one would normally associate with a large pseudotumour, and it was not “severe”: there were no granuloma or germinal centres, and very little perivascular cuffing. The macrophage pattern was very different from that seen in the cases of Mr Haley and Mrs Emery. It was consistent with a mild immunological reaction to metal debris that was insufficiently severe to cause any tissue damage, and/or chronic inflammation consistent with the underlying osteoarthritis. If what Mr Nargol saw was indeed a “*very large pseudotumour*” which was responsible for the damage to the abductors then, taking into account all the relevant expert evidence at trial, it is more likely than not that the histological findings would have been of a similar nature to those seen in the cases of Mr Haley and Mrs Emery. The histopathology is not in keeping with what Mr Nargol recorded in his note. It is impossible to reconcile them. Everything else the histopathologist recorded was consistent with what the experts saw on the slides (apart from Professor Freemont’s finding of tissue necrosis).
676. It is important to remember that the difference between a “pseudotumour” and a purely extra-articular collection, which may otherwise look identical, is that the former connects with the hip. There was no evidence of any such connection and it seems rather unlikely that a connection that was so thin as to be invisible on scans to three eminent radiologists who were looking for it (including Dr Raju) would lead to a pseudotumour of those dimensions. A failed abductor repair would account for the presence of some metal debris in the bursa (though the histology indicated very little visible metal even under a very high-powered microscope).
677. Above all, Mr Nargol was expecting to find a pseudotumour; Dr Raju’s reference to an “abundant” joint effusion had conditioned that expectation, and therefore he saw what he was expecting to see. There is force in DePuy’s observation that Mr Nargol’s attitude to MoM hips, as demonstrated in his evidence, made him prone to characterise any fluid collection in the area of the greater trochanter as a pseudotumour, regardless of its size or features. There is no evidence that he even considered the possibility of bursitis, or that the phenomenon he observed might have been a sign of a chronic inflammatory process brought about by attempts to heal a failed abductor repair. As is shown by his attitude when giving evidence, once Mr Nargol has made his mind up about something, he will not accept any other possibility.



678. Mr Oppenheim raised a number of factors which he submitted supported a correct diagnosis of ARMD. He rightly pointed out that even if there was a failed Hardinge repair, that does not mean there was no pseudotumour. Indeed, he submitted, the fluid from the joint could have reached the greater trochanter through the site of the failed repair, which means the absence of a visible connection on the MRI images is less significant than might otherwise be the case. Whilst that is theoretically possible, there is still no evidence that there was any effusion from the joint itself.
679. Some of the factors on which Mr Oppenheim relied take Mrs Stalker's case nowhere. Nothing can be deduced from the fact that Mrs Stalker had slightly elevated (but fluctuating) levels of blood metal ions prior to the revision, and they came down afterwards; that is what one would expect of any patient who had an implant with a MoM articulation that was replaced by other materials and therefore was no longer in a position to shed metal debris. The blood ion levels are not a diagnostic factor, but a means of screening those patients who may be at risk. I have rejected Mr Nargol's evidence that the aspirate taken from the hip joint was opaque. Even if it were, nothing can be deduced from that or from its colour. Mr Oppenheim referred to Mr Kim's acceptance that the weakness of the external rotators could not be ascribed to a failed Hardinge repair – but whilst ARMD was one possible explanation for that feature, nobody suggested that it was the only explanation. She had weak external rotators even after the revision surgery but there was nothing in the operation note that indicated that any muscles other than the abductors had been damaged.
680. Mr Oppenheim also relied on the size of the tissue samples excised and sent off for analysis. Professor Kay had referred in his evidence to the difficulties that surgeons encountered in excising the contents of pseudotumours; Mr Oppenheim submitted that it would be odd, to say the least, for a surgeon to excise so much tissue that his patient required a brace, if what he was really faced with was a case of chronic tendinopathy following a failed Hardinge repair. Although that submission has a superficial attraction, there are a number of answers to it. If the MRI image was indeed of an area of "boggy tissue" as Dr Wilson thought it might be, that might explain the extent of the excision. However even if it was purely a fluid collection, but he was working on the assumption that he was dealing with a pseudotumour, Mr Nargol would expect to find soft tissue damage in that area. What he found were signs of chronic inflammation, and thus it is unsurprising that he tried to ensure that he had cleared out anything that he believed might get progressively worse and cause further problems in the tissues around the joint even after he had swapped the head and liner.
681. In my judgment, the absence of necrosis in the tissue is the fatal piece of evidence and it is something the claimants cannot get around; in samples that large, if there were any necrotic tissue, it would have been seen. Professor Kay came up with no convincing explanation for the absence of any evidence of soft tissue damage, and his opinion that there was ARMD depended very much on the existence of the pseudotumour.
682. It is not for DePuy to establish an alternative medical explanation for the findings at the time of the revision operation. The burden is on the claimant to establish, on the balance of probabilities, that she suffered from ARMD, and in my judgment, standing back and considering the whole of the evidence, she has failed to do so. Had I been persuaded otherwise, the evidence indicates that Mrs Stalker has made a very good

recovery from her revision surgery, and that the problems she has experienced since then are referable to other factors.

## **5.5 PATRICIA GARRATT**

683. Mrs Garratt was born on 30 June 1933. She gave her evidence via video link. Unlike the other five lead claimants, Mrs Garratt's Pinnacle prosthesis was not a primary implant, it was put in as a revision. That gives rise to insuperable problems in terms of causation, even if Mrs Garratt could prove that she did develop ARMD and her Pinnacle prosthesis was revised because of it. There was no evidence from which the Court would be able to infer anything about the likely survivorship of an implant that was used in a revision procedure. The claimants' case depended on a comparison of the risks of failure of *primary* implants within the first 10 years after implantation, and the experts were agreed that revisions were generally associated with worse outcomes than primary surgery, though the outcomes in an individual case depended on a variety of factors.
684. Mrs Garratt had a conventional MoP total hip prosthesis implanted on her right side on 8 May 1997, when she was 63 years old. On 1 December 2004, she had a CoP total hip prosthesis implanted on her left side. On 27 March 2008, the right prosthesis was revised, and she was provided with a Pinnacle MoM prosthesis on the right side. She was then aged 74. On 22 March 2011, only three years after implantation, the Pinnacle MoM prosthesis was revised. The head and liner were exchanged.
685. Both the right revisions were carried out by Mr Clayton Marsh, who gave evidence at trial. Mr Marsh understandably had no independent recollection of the details of the events relating to Mrs Garratt and was heavily reliant upon the medical records to refresh his memory. Whilst I believe he was doing his best to help the Court, he was prone to speculate without making it clear that this was what he was doing, and in consequence I had to treat his evidence with a considerable degree of caution. He was forced to acknowledge that several parts of his witness statement were incorrect, or pure supposition, and abandoned them under cross-examination.
686. Mrs Garratt, too, had no independent recollection of events, but unlike Mr Marsh she refrained from filling in the gaps with what she thought must have happened. She told Mr Owen that she had difficulty in remembering the chronology of her symptoms, which is hardly surprising. I concluded that the medical records were the most reliable source of evidence in this case.
687. Mrs Garratt had degenerative changes in her spine that were noted as early as 1987, and osteoarthritis in both knees was noted in 1992. She experienced pain in her hips since the early 1990s. An X-ray of both hips in June 1992 showed moderate degenerative changes of osteoarthritis affecting the right hip joint, although the left hip joint remained relatively normal at that time. Nothing was done at that point, possibly because it was felt that she was too young. In February 1997, the pain in her hips was sufficiently bad for her to consult her GP. He noted that clinically they were not too bad, as she had a reasonable range of movement, but referred her to a consultant orthopaedic surgeon, Mr Smibert.
688. Mr Smibert discussed the options available with Mrs Garratt at some length. They were either to proceed to surgery straight away, or to put up with the pain, though he

said he would need to keep an eye on the bone stock and if she were shown to be losing any bone they would need to proceed to a hip replacement. He explained the procedure for a hip replacement including the potential complications of infection, loosening, DVT, and dislocation. He also said that he could not guarantee that her leg lengths would be the same.

689. Mrs Garratt opted for the surgery, and a right total hip arthroplasty was carried out by Mr Smibert in May 1997 using the Hardinge approach. He noted that the procedure went uneventfully save for some oozing. He also noted that her left hip had deteriorated since he last saw her. Histology tests for rheumatoid arthritis proved inconclusive. At a review seven weeks later, she appeared to be making excellent progress and mobilising well with a good range of movement. Mr Smibert informed the GP that because of the early changes in the opposite hip he would need to keep Mrs Garratt under review and see her in a year's time with X-rays of both hips, or earlier if she had any problems.
690. In October 1998, Mrs Garratt was reported by Mr Smibert as coping with the left hip. She was experiencing some aching in her lower back, but he did not think that was related to the hip. Four years after the total hip replacement, in 2001, Mrs Garratt was reported as being delighted with her hip. However, in 2003, when she consulted a cardiologist for an unrelated problem, he noted that her exercise was currently limited by painful hips. On 10 August 2004 she saw the GP because of left hip pain, and an X-ray confirmed evidence of degenerative change in that hip with mild/moderate loss of the joint space and a little marginal osteophyte formation. She was subsequently referred to Ms Muirhead Allwood at the London Hip Unit, with a view to having a left total hip replacement. Mrs Garratt had been carrying out some research and was interested in a new procedure that Ms Muirhead Allwood was carrying out; in oral evidence Mrs Garratt said she may have read an article about this in a newspaper.
691. In November 2004 Ms Muirhead Allwood reported to the GP that Mrs Garratt's right hip looked satisfactory, apart from the possibility of a slight socket demarcation, but that the left hip was progressively painful, and she was having increasing problems over the trochanter and groin. Increased pain in the right hip over the last few weeks was likely to be due to compensation for the left hip pain by walking more on the right side. The left total hip arthroplasty was carried out on 1 December 2004 using what appears to have been a new design of Zimmer CoP prosthesis, and Mrs Garratt appears to have had an uneventful recovery.
692. There were no further problems reported with her hips until 19 February 2008 when she went to see the GP complaining of recent right hip discomfort and pain. She said she wanted to see a private consultant again, and was referred to Mr Marsh at Somerset Nuffield hospital. Mr Marsh's letter to the GP following his consultation described Mrs Garratt as a very interesting example of the evolution of hip arthroplasty. He stated that her right Charnley had become painful in the last 12 months or so, and showed typical signs of loosening of the acetabular component, which was to be expected at 10 years plus. In his evidence Mr Marsh agreed that radiological imaging would have shown gaps between the bone and the prosthesis and that the "cup effectively rattles within". He said that most studies say that 10% to 15% of such implants are loose at ten years, and agreed that it was not a great surprise when problems arose with a Charnley hip after about ten years, especially with an active patient (as Mrs Garratt was).

693. Mr Marsh's advice was that the prosthesis should be revised. He reassured Mrs Garratt that it was technically a very straightforward matter to convert to an appropriate modern cementless design with a hard bearing (i.e. one where both the head and liner were made of hard materials). She was keen to proceed with this. Mr Marsh explained in his oral evidence that the Charnley stem had to be changed because it had a 22mm head and could not be retained for modern designs. He offered Mrs Garratt a Pinnacle hip with a 36mm femoral head and liner in a 54mm cup, because of its stabilising qualities. He viewed this as important because the incidence of dislocation in revision surgery is much higher than in primary surgery, and the large head gave a much more stable platform. He said that if he had wanted to use a material other than metal, he would not have been able to use such a large head. He considered that Mrs Garratt was fit for her age and that she could well live for another 15 years.
694. When she was assessed in preparation for her hip revision surgery, Mrs Garratt was noted as doing very little exercise at present and she could only walk for about 10 minutes two or three times a week. She was having to use a stick while she was out of the house. Mrs Garratt does not drive and therefore she used to walk everywhere.
695. On 27 March 2008, Mrs Garratt underwent the revision of her right hip prosthesis. Mr Marsh had to extract the first prosthesis completely as well as excising the cement. Aseptic loosening of a primary Charnley hip usually demanded complete replacement of the prosthesis, which was more complex than a simple head and liner exchange. There was a trochanteric fracture during the operation, which Mr Marsh repaired using two Dall-Miles cables to hold the bone back in place. This was a normal incidental risk of removing the original prosthesis. Mr Marsh did not alarm Mrs Garratt by telling her about this afterwards; he said he would not have done so unless he needed to carry out a massive reconstruction procedure. He fitted a KAR stem because this was revision surgery. He explained that it was the middle stem of the Pinnacle range and extended beyond the area of cement and into virgin bone where you can get good ingrow. The records of her follow-up indicate that she made a good recovery with satisfactory X-rays.
696. In February 2009 Mrs Garratt suddenly developed a pain in her other hip and went to see her GP about it, but her revised right side was described at that time as being very comfortable. In August 2009 she saw her GP with ill-defined pelvic pain which she said was more on the right side in the left, but by the time she saw Mr Marsh again in November her symptoms on the right side had settled considerably, and her revised right hip was said to be more comfortable than her left primary hip replacement. Mr Marsh reported to the GP that she was now walking well and that the discomfort in her revised right hip had settled. Mr Marsh noted she had some "*low grade discomfort*" in her left hip which needed to be watched.
697. In April 2010 Mrs Garratt was complaining to her GP about back pain which she associated with the revision surgery on her right hip. Mr Marsh saw her on 26 April. He described a sudden onset of pain in her right leg as being associated with back pain and pain down into the legs. He noted that her walking distance decreased due to tiredness in the legs, which appeared to be of spinal origin. On examination, her right total hip arthroplasty was freely mobile, she had a negative Trendelenburg test, there was no pain on walking and no significant features on X-rays. Mr Marsh thought that was all encouraging for a revision. Mr Marsh reported to the GP that the symptoms

sounded very much like low-grade spinal stenosis. Mrs Garratt did not want that investigated at that stage, but his next plan would be to have an MRI of the spine if the pain persisted. However, in the event, the pain appears to have settled.

698. On 16 June 2010, Mrs Garratt was sent a standard letter from the Nuffield hospital prompted by the issue of the MHRA guidance about MoM total hip arthroplasties. It stated that a small number of patients had developed problems after MoM surgery that could present with pain and reduced mobility. The MHRA had given guidance about such hip replacements which might impact on the patient's follow-up and treatment. Where there was pain or symptoms of potential problems such as reduced mobility the patient should seek specialist assessment. If they were concerned or had not seen a consultant for over a year they should seek a further assessment.
699. Mrs Garratt went to see her GP on 25 June, and drew the letter to his attention, but as she had seen Mr Marsh in April and was due for a follow-up appointment on 25 October, he took no action other than to prescribe some further painkillers. When Mrs Garratt saw Mr Marsh in October 2010 he noted that she was still having some "low-grade discomfort" in her hip. The previous neurological type pains had disappeared, but Mr Marsh said he was slightly worried because she had a MoM bearing surface. I am satisfied that his contemporaneous notes and letter to the GP are more likely to be a reliable record of Mrs Garratt's symptoms at that time than what she said in her witness statement, which was that she was starting to have trouble with pain, her mobility was greatly restricted again, she was getting concerned about her right hip due to the discomfort and she was having to take taxis more often. That description seems more in keeping with the symptoms she was experiencing immediately prior to the first revision by Mr Marsh in which the Pinnacle MoM hip prosthesis was implanted.
700. Mr Marsh agreed that low grade discomfort would not necessarily have rung any alarm bells in a non-MoM revision prosthesis unless there were signs of loosening on radiographs or bone scans. Low grade discomfort is not uncommon for someone who has had a revision replacement, because of the extent of scar tissue formation. I am satisfied that Mrs Garratt's previous symptoms had settled, she was not experiencing any significant pain, and it was Mr Marsh who decided to carry out further investigations, solely because she had a MoM prosthesis. He told her GP he was going to send her for some blood ion investigations and arrange an ultrasound of the hip to look for a fluid collection.
701. In his oral evidence Mr Marsh explained that his concerns were due to his personal experience at that stage. He had stopped using MoM implants in 2009 following a conference in the USA, where American surgeons had presented papers about premature failure rates due to reactions to metal debris. This may well have been the same conference that had a similar influence on Mr Herlekar, who operated on Mr Woods. Mr Marsh was aware that concerns were being expressed in the orthopaedic community that, besides early failure, there may be systemic problems due to the metal debris. By that time, he had a number of other patients who had had severe adverse effects from MoM (he did not say whether these were resurfacings or total hip arthroplasties or a mixture of the two).
702. Mrs Garratt's blood ion test results showed mildly elevated levels of chromium and cobalt for a patient with a MoM implant – the equivalent of Co 7.21ppb and Cr 8.55

ppb. Mr Marsh agreed in cross-examination that the levels were above those at which there was no cause for concern, but well below the levels that were indicated as being likely to be associated with ARMD (which he referred to as ALVAL), though he said, quite correctly, that a patient could have low ions and still experience a great deal of trouble. He very fairly agreed, when it was put to him, that he misunderstood the results at the time, because they were measured in nanomols and microgrammes and not parts per billion, and therefore at the time he thought they showed much higher levels. They were, however, slightly above the MHRA screening threshold.

703. The ultrasound was carried out on 24 December 2010 by Dr David Cooke, a Consultant Radiologist and senior lecturer at the University of Bristol. Mr Marsh described Dr Cooke as “a very skilled radiologist”. His report indicated a fluid collection around the right greater trochanter measuring 36×25×5 mm. The fluid collection had tended to wrap around the greater trochanter. It did not say there was any communication with the hip joint and there was no report of any joint effusion.
704. Mr Marsh considered that this fluid collection was consistent with an adverse reaction to MoM components; the idea that it might have been a bursitis does not appear to have crossed his mind. Mr Marsh’s explanation for failing to ask Dr Cooke to look again for any communication was that he was sure that he would have done “*as assiduous an investigation as he could, given the equipment available to him at that time*”. However, it became clear that Mr Marsh would not have asked Dr Cooke to look for a communication anyway, because he did not seem to appreciate that an extra-articular collection can only be classified as a pseudotumour if it communicates with the hip joint.
705. Mr Marsh accepted that if a patient with a MoM hip was in pain, that would be enough for him to revise the hip. In this respect, his attitude was similar to that of Mr Dunlop and Mr Herlekar. In his letter to the GP he said: “*my experience with this situation is steadily developing and my advice is that Mrs Garratt should seriously consider having the metal/metal bearings changed to a metal/polyethylene. Unfortunately, this does mean exposing the prosthesis surgically, but I have reassured her that the femoral and acetabulum components per se do not have to be dismantled, and to just simply change the lining in the acetabular component and the femoral head. This will be very successful in relieving this current situation. She quite understandably wishes to think about this and I would be very happy to see her again any time if she wishes to proceed.*”
706. Mrs Garratt did wish to proceed – her evidence was that she “*didn’t want to wait around for it to get worse*”. There was no contemporaneous evidence in her medical records that her symptoms had declined since Mr Marsh last saw her in October 2010, and it was clear from her oral evidence that she could not recall that they had. Although there was no reference to it in his witness statement, Mr Marsh volunteered in his oral evidence that she may have spoken to him on the telephone and told him her symptoms had got worse, but there is no note of any such conversation and it transpired that he was speculating rather than recollecting any such communication.
707. Thus the reasons for the re-revision were that the patient had a MoM prosthesis, she experienced mild discomfort, there was imaging showing a small collection round the trochanter and her blood ion levels were mildly elevated, though Mr Marsh thought they were higher than they actually were. Mr Smith, the claimant’s orthopaedic expert

in this case, agreed that those factors would not justify revision. Mr Marsh adopted a very different course to the one he adopted when she experienced problems with her CoP hip on the other side. It appears that, apart from his concerns about MoM, he was influenced by the relative ease of swapping the head and liner. He said that this was a modular system, so the change was simple.

708. Mr Marsh's notes of the operation on 22 March 2011 indicate that he immediately encountered a large collection of bloodied fluid in and around the hip arthroplasty. There is no specific mention of the fluid collection round the trochanter that Dr Cooke had found on the ultrasound. Mr Marsh described the lining membrane of the cavity as "very typical" of this situation. He sent a sample for histology. He replaced the head and the liner as planned with a different metal head and a polyethylene liner, and said that this gave a stable arthroplasty.
709. During the re-revision surgery Mr Marsh removed one of the Dall-Miles cables that had been fitted to repair the fractured trochanter. In his witness statement he said he could not recall whether he removed one cable or both, but explained that he would have removed them because "*the fracture had healed and the cables were redundant*". However, in cross-examination when he was asked about this he then said that he removed one cable and left the other behind because that was his standard practice. When asked why he would remove a cable that was doing no harm, he said that it may have blocked the surgical approach to the femoral trunnion. Mr Owen was adamant that this explanation made no sense. It became clear, and eventually Mr Marsh accepted, that he had no independent recollection of what he did in Mrs Garratt's case and he could only speculate as to why he would have removed only one cable.
710. The histopathology report dated 1 April 2011 stated that the features were those of inflammation and hyperplasia with fibrosis. The typical lymphocytic reaction seen in ALVAL was not seen and the histopathologist said "*I think the likelihood that this condition is present is low. There is no evidence of a neoplastic process.*"
711. In May 2011 Mrs Garratt had a fall whilst on a bus, landing on her sacrum. Her GP saw her on 5 May and noted that she was now mobilising with sticks. An orthopaedic consultant named Mr Squires saw her on 9 May and noted that Mrs Garratt was making an excellent recovery from her surgery until she had the fall.
712. In August 2011 Mr Marsh retired from NHS practice. He stopped operating the following March and he retired altogether in July 2012. This explains why Mrs Garratt had no further consultations with him. The next specialist she saw was Mrs Acton, a specialist musculoskeletal practitioner, in December 2011. She recorded that Mrs Garratt was having problems with her hip, and that there was weakness on external rotation and abduction. X-rays revealed that the trochanter was avulsed at the attachment of the gluteus medius muscles, and Mrs Acton considered that to be the probable cause of her weakness. In all probability, the fracture was caused by the fall on the bus; that was Mr Marsh's opinion, and if there had been a fracture during re-revision surgery I would expect him to have made some note of it.
713. There are sporadic references in the records to hip pain in 2012 and 2013, but this appears to have been controlled by analgesia. Mrs Garratt was not referred back to the

consultant during this period, as she would have been if there were any significant cause for concern.

714. On 19 May 2014, Mrs Garratt attended her GP complaining of “new” pain in both hips, though it was said to have settled by the time she saw the GP. That suggests that she had been relatively pain-free up to then. There is then no relevant entry in the medical records relating to her hips until 6 June 2016, when her GP saw her and noted that she was unable to lie on her side “*because of her hip replacements*”. There was a similar entry on 2 December 2016 when the GP noted that Mrs Garratt was normally fit, well, independent and fully mobile but had difficulty sleeping “*due to bilateral hip operations*”.
715. In June 2017 Mrs Garratt had an onset of pain in her left hip (i.e, the non-Pinnacle prosthesis) which caused a “spasm” as she was making the bed, though in her oral evidence she clarified the pain was in her groin. She had to drag herself to the telephone, as she was unable to bear any weight on that side. Her GP wrote a letter on 11 October 2017 explaining that she was now housebound as a result of her pain, which is why she gave her evidence by video link.
716. DePuy suggested that the most obvious explanation for Mrs Garratt’s presentation at re-revision was that one of the Dall-Miles cables used to fix the previous fracture was causing irritation and inflammation of the soft tissues in the region of the right greater trochanter. That was Mr Owen’s view. Mr Smith disagreed, but he did accept that the cable could rub and irritate overlaying soft tissue and that there was a potential for fretting at the bone cable interface and at the crimpsleeve cable interface which could cause problems.
717. In my judgment, that was the most obvious explanation for Mr Marsh removing only one of the cables when he carried out the re-revision. Whilst he speculated in the witness box about why he did it, the truth is that he could not remember. Mr Owen described the cables as braided, like multifilament steel rope. They are stronger than Charnley wires, which could break more easily, and bone grows into them, making them difficult to remove. The cables here were well fixed, and therefore a surgeon would only remove them for a good reason. Mr Owen said the only reason to remove them would be if they were loose or broken or causing irritation or if the hip was infected, otherwise the risks of taking them out outweighed the advantages. The only one of those reasons that arose in Mrs Garratt’s case was irritation, and that is consistent with the local chronic inflammation and the blood in the fluid being caused by a haematoma.
718. If the reason for removing the cables was that they were no longer needed, logically Mr Marsh would have removed them both, not just the one nearest the greater trochanter. The only reason he gave for removing only one (apart from suggesting for the first time in oral evidence that was his normal practice, which does not explain the reason for that practice, or why he did not mention it in his witness statement) was that it had a potential for obstruction during surgery. That was something which he also volunteered for the first time in his oral evidence. It was not an explanation advanced by Mr Smith, and Mr Owen’s evidence was that it made no sense. It would make far more sense to remove a source of irritation that was no longer serving a useful purpose.



719. Dr Ostlere and Mr Wilson agreed that the ultrasound images show no joint effusion, and no communication between the extra-articular collection and the hip joint. The size of the collection was in keeping with a bursitis. Dr Cooke made no suggestion that it was a pseudotumour. He was a very experienced radiologist and therefore was likely to have checked for any connection. Contrary to Dr Ostlere's supposition that Dr Cooke had not recorded any images of the neck of the prosthesis, and therefore it was possible to infer that he had not checked this area for a fluid collection, Dr Wilson's subsequent checks established that there were such images and there was no evidence of fluid on them. Dr Wilson's evidence, which I accept, was that there was an image of the hip and that it showed no evidence of any fluid in the hip joint.
720. The surgical findings are inconsistent with ARMD. There was no soft tissue damage noted and Mr Marsh confirmed that he would have noted it if he had seen it. The collection of bloody fluid that Mr Marsh noted he encountered immediately is likely to have been the collection observed on the ultrasound, given that it would be the first thing he would hit on the surgical approach that he took. Although the operation note refers to the collection being "*in and around the hip*", Mr Marsh did not generally distinguish between the hip joint and the trochanter when he gave his evidence. There is nothing recorded in the operation note to suggest that the collection tracked down to the hip joint itself. Unless there was some evidence of a connection, it cannot be inferred from the fact that there was a collection in that area that there was ARMD, as Mr Smith accepted. When Mr Marsh sent off the tissue samples to histopathology he wrote "*? early pseudotumour*". That is inconsistent with his encountering an actual pseudotumour. He was plainly unsure whether the fluid he found was the sign of a pseudotumour amassing in that area (Mr Smith said that he, Mr Smith, would not have used the expression "early" but that appears to me to be what Mr Marsh was driving at.)
721. The histological evidence is inconsistent with ARMD. There is no reference to any tissue necrosis. Professor Nelson's evidence, which I accept in preference to that of Professor Freemont, is that no appreciable necrosis is present on the slides. Professor Freemont alleged that there was "*superficial*" necrosis. Professor Nelson provided an image of the surface of the joint which shows fibrin accumulating on the surface, but no evidence of any necrosis. Professor Freemont accepted that this was the case on that image. When Professor Freemont tried to rely on other (low magnification) images as showing necrosis, Professor Nelson demonstrated convincingly that what Professor Freemont described as a tissue cell in a hole was in fact a macrophage in the fibrin, which is not an uncommon finding.
722. The findings are inconsistent with an ALVAL reaction and the author of the contemporaneous report indicated that the chances that this condition was present were low. Whilst some lymphocytes were present, the predominant lymphoid cells that were seen were plasma cells. Professor Freemont, undeterred, espoused a novel theory that ALVAL can be diagnosed where the predominant finding is plasma cell cuffing. As DePuy pointed out, ALVAL is an acronym for Aseptic Lymphocyte-dominated Vasculitis-Associated Lesion. Although Professor Athanasou published a paper about ALVAL in 2016 which describes a "*pronounced often heavy perivascular lymphocyte (and plasma cell) reaction*", there is no support in that paper for a diagnosis of ALVAL where the predominant finding is plasma cell cuffing. Professor Athanasou's description is in keeping with Dr Willert's seminal paper on ALVAL

which describes the typical finding as a “*distinct lymphocytic infiltration, sometimes accompanied by plasma cells.*” In this case there was very little, if any, perivascular lymphocytic cuffing and it was certainly not predominant. I accept Professor Nelson’s evidence that there was no ALVAL. It is consistent with the contemporaneous histological findings.

723. There was a dispute between Professor Nelson and Professor Freemont about whether metal debris could be seen on the slides. Professor Nelson gave clear and cogent reasons why the visible debris was likely to be from the bone cement removed by Mr Marsh when he carried out the initial revision operation. The dark particles were found around the outline of a hole left where bone cement would have been. They were unlikely to have been metal oxide because they were too large and not opaque. Professor Freemont’s evidence about this was somewhat confused. When the images were subject to higher magnification, he accepted that some of the particles might be bone cement, but pointed to others as being too small, an explanation given for the first time in his oral evidence. I prefer the evidence of Professor Nelson.
724. There was evidence of a macrophage response, but that could have been a response to the bone cement residues or to other material such as hemosiderin, a breakdown product of blood, which was present on the slides, a not uncommon finding after a total hip arthroplasty. Taken as a whole, the histopathology in Mrs Garratt’s case does not support a clinical diagnosis of ARMD. There was no other evidence to support that finding on the balance of probabilities, and therefore I conclude that Mrs Garratt did not have ARMD. In any event, there is no evidence that her outcome from the re-revision surgery was worse than it would have been if her implant had been re-revised for some other reason.

## **5.6 PETER WOODS**

725. Mr Woods is the youngest of the six lead claimants; he was born on 26 October 1967. He had a left Pinnacle MOM prosthesis on 6 November 2007, at the age of 40. He underwent a revision on 31 January 2014. The head and liner were exchanged for ceramic components; the cup and stem remained.
726. When he was a teenager, Mr Woods played a lot of football, but he began to experience pain in his left groin, particularly towards the end of a game. Over the years the pain got worse, and so he went to see his GP in the autumn of 1998 to investigate the cause. The GP carried out an investigation and suspected he was suffering from a condition known as Gilmore’s groin. He referred Mr Woods to a consultant surgeon at Furness General Hospital, Mr Allan. The consultant confirmed the diagnosis, describing him as presenting with “*a typical history of left Gilmore’s groin*” and remarked that he was very fit in other respects. Mr Woods underwent a repair procedure in March 1999.
727. Whilst the operation improved matters, the groin pain returned about a year later, and further investigations were carried out. In April 2001 his GP recorded that his left groin pain was much better now, but he now had pain on the right which was almost as bad as on the left. Eventually, in around March 2003, after the review of his latest X-rays by Dr O’Connor, a consultant musculoskeletal radiologist, he was diagnosed with osteoarthritis in both hips, the left one being worse than the right. His GP referred him to a consultant orthopaedic surgeon, Mr Hodgkinson, at Wrightington

Hospital. The referral letter stated that his left hip was particularly stiff, with decreased internal rotation, and that he had a very poor FABER test.

728. Mr Hemmady, an arthroplasty fellow, examined Mr Woods in clinic on 18 August 2003. He recorded that the pain in the left groin was situated anteriorly with some radiation along the medial aspect of the thigh. Mr Woods was getting a dull aching sensation in the groin towards the end of the day, but experienced no functional disability at that time. He told Mr Hemmady that he did not take painkillers, and that he could easily walk for about 4 to 5 miles without much difficulty. Mr Hemmady decided to discuss the matter with Mr Hodgkinson on the latter's return from his holidays, because although it appeared that Mr Woods had early osteoarthritic changes, it was rather difficult to be sure about whether the source of his pain was intra- or extra-articular. Mr Hodgkinson's opinion was that the symptoms were mainly due to early degenerative changes in Mr Woods' hip joint. Given the level of those symptoms Mr Hemmady told Mr Woods' GP it would be best to leave things well alone for the time being, but that they would like to review him in six months' time.
729. On 19 May 2004, Mr Hemmady noted that the equivocal picture remained much the same, in that Mr Woods had intermittent symptoms from his left hip, but he was working full time and sleeping well at night. They had a long discussion about the options available at that stage and on balance, in view of his relative youth, it was mutually agreed to leave things well alone for the time being, and to review him in two years' time. Mr Woods was told that it was sensible to postpone having a total hip arthroplasty for as long as possible because hip prostheses need to be replaced from time to time, and that since he was only in his mid-thirties, it was rather early to be starting the process.
730. The next occasion on which Mr Woods was examined by a consultant was in October 2006. By then he had moved away from Mr Hodgkinson's team to Mr Garg, who was based in Lancaster Royal Infirmary, which was nearer to Mr Woods' home. Mr Garg noted that Mr Woods had had a painful left hip now for some 10 years which had been getting worse in the very recent past. Despite this, Mr Woods was still able to manage his work as a warehouse manager. He was prescribed analgesics and told that if things became unbearable they would think about hip resurfacing or a MoM total hip replacement. Mr Garg explained that a conventional hip prosthesis could fail very quickly in a young and active man. but said that there were some new types of prosthesis that might have a longer life. Mr Garg suggested that if he opted for a total hip arthroplasty Mr Woods might want to try one of the newer types in the hope that they would serve him better.
731. Mr Woods' condition continued to deteriorate as time went by. At the time of his next review, which was carried out by one of Mr Garg's associates, a Mr Sinha, on 26 March 2007, he had persistent pain in the left groin and occasional pain in his back, though Mr Sinha thought the back pain was possibly unrelated to his hip. An X-ray of the pelvis confirmed "quite significant arthritic changes, especially on the left side." Mr Sinha's note of the consultation states that he had explained the options of resurfacing or a MoM total hip arthroplasty. It also states that he explained the risks "including infection, revision and blood clots etc. He may be on crutches for about 6 weeks." Mr Sinha noted that he thought that Mr Woods had made up his mind to go

for surgery, but he wanted to come back in six months, and that he would possibly be listed for a hip arthroplasty then.

732. At the next review, on 17 September 2007, Mr Garg noted that Mr Woods' left hip pain had become so bad that it was keeping him awake at night and affecting his walking as well as his work. Clinically and radiologically he now had gross osteoarthritis of the left hip. There was a discussion about the pros and cons of a total hip arthroplasty, which included a discussion of the potential advantages of a MoM prosthesis over a conventional prosthesis. Mr Woods was told that with a MoM prosthesis he would get better function, a better range of movement, and that the prosthesis would be more durable, all of which would be beneficial for a young and active man like himself. Mr Garg did not put a timescale on how long the MoM prosthesis would last but indicated that it would last longer than a conventional prosthesis. Mr Woods signed the form consenting to surgery on the same day.
733. Mr Woods had little recollection of the risks that were pointed out to him, though he accepted that he would have been told of them when he signed his consent form. He did not remember the process of signing the consent form, and the only risks he was now able to remember being mentioned to him specifically were infection and dislocation. However, it is plain he was told about the risk of revision six months earlier, and it is unlikely that Mr Garg would have omitted that risk from "all the pros and cons" that his letter to the GP stated he had explained to the patient.
734. The total hip arthroplasty was carried out by Mr Garg on 6 November 2007 using a Corail stem with a 36 mm head. He took the Hardinge approach. Mr Woods made a good recovery from the surgery and at his check-up on 17 November, was reported to be very pleased with the results. Mr Woods made a deliberate decision not to go back to impact sports following the operation, but he was able to carry out a full range of normal activities. At his six-month review on 14 April 2008, he was described as having made "*excellent progress*", with satisfactory x rays. He was Trendelenburg negative and fully independently mobile, back at work and otherwise comfortable. Likewise, when Mr Garg saw him at his 12-month review in November 2008 he described him as "completely symptom free". Mr Garg noted that X-rays of the left hip showed some heterotrophic calcification around the greater trochanter area, but this was not causing any symptoms. He may not even have mentioned this to Mr Woods, who had no recollection of it.
735. Mr Woods did experience intermittent lower back pain. In January, and again in February 2008 he went to his GP complaining of pain in his lower back which he reported as having existed prior to his hip replacement operation and which had gradually worsened. There is no further reference to back pain in his medical records until June 2010 when he complained of pain after working in the garden, and then again in May 2012, when the back pain appears to have made it difficult for him to stand, although when he came to give his oral evidence, Mr Woods had no direct recollection of that episode. On each occasion, the GP treated the pain with analgesics, and appears to have diagnosed muscle strain. Mr Woods was never off work with back pain.
736. In June 2012, Mr Woods was reviewed in the special clinic set up at the Royal Lancaster Infirmary for review of MoM hip replacements. He went to the clinic in response to a standard letter that was sent out to all patients who had undergone a total

hip arthroplasty in the department, using a MoM prosthesis. He was examined on 30 June 2012 by Mr Herlekar, one of the two consultant orthopaedic surgeons carrying out reviews in that clinic, who gave evidence at trial. Mr Woods' X-rays appeared to be fine, apart from the heterotopic ossification, which Mr Herlekar accepted was something that one would see from time to time, and which would not always cause discomfort or pain. The heterotopic ossification did not appear to be having any ill effect on Mr Woods at that time. He had no symptoms in his left hip and remained very pleased with his left MoM hip replacement.

737. Mr Herlekar explained to Mr Woods that some concerns had been raised by the MHRA about 36mm MoM hip prostheses, and told him that if anything changed, he should contact the clinic immediately. He ordered blood ion tests as part of a screening process. When they came back on 17 July 2012, they indicated very low levels of cobalt and chromium, way below the levels that might be indicative of cause for concern in a patient with a MoM prosthesis, even after taking into account the recently updated MHRA guidance published in June 2012. Those results were sufficiently reassuring that Mr Herlekar decided there was no reason for concern. Mr Herlekar's detailed letter to Mr Woods' GP following the consultation makes no mention of Mr Woods' long history of pain in the groin prior to the primary operation.
738. Mr Woods was next referred to the clinic by his GP in January 2013 after developing some "non-specific" pain in the lateral aspect of his left thigh over the previous three weeks. He saw a member of Mr Herlekar's team, Mr Yousuf, and then Mr Herlekar himself on 28 January 2013. Mr Woods denied experiencing any pain around the trochanteric region and any significant pain in his groin area on the left side. Mr Yousuf and Mr Herlekar shared the view that the symptoms Mr Woods described were attributable to referred pain from his back. Mr Herlekar explained in his oral evidence that there is a specific pattern to such pain, which radiates down to the buttocks around the back of the thigh, and that where there is such a presentation, the patient is told that it is a back-related problem. Mr Yousuf carried out a thorough neurological examination which indicated no neurological abnormality. There was also no trochanteric tenderness and a good range of hip movement. X rays showed a satisfactory placement of the implant.
739. Mr Woods came back to the clinic on 24 September 2013. On that occasion he was recorded by Mr Herlekar as having pain and discomfort in the groin, yet on clinical examination there was no grossly abnormal finding. It was the evidence of both orthopaedic experts that the pain could have been part of the long history of groin problems that pre-dated the primary surgery, or that it could have been associated with his recurring problems of back pain. It was even possibly just a case of periodical and fluctuating pain that one might expect to experience from time to time after primary surgery.
740. Despite this, Mr Herlekar decided to investigate for ALVAL (as he termed ARMD) and ordered an MRI scan, blood ion tests and aspiration of the left hip under X ray control. His letter to the GP states that "*these hips are known to develop ALVAL type of reaction at around 5 to 6 years. I am hence investigating this further by the above-mentioned tests. I have explained to Mr Woods that he is very likely heading for a revision of this hip*". In cross-examination, Mr Herlekar frankly admitted that his mindset at that time was that if a patient had a MoM hip and some unexplained pain,

it was very likely that he was going to end up revising it. If it had been a different type of articulation, the idea of revision would not even have crossed his mind.

741. He explained his thinking on the basis that by that time, he had had to carry out revisions of around 20 hip prostheses, some of which were 36mm MoM, in which he had seen what he described as “*horrendous destruction*”, some of which were difficult to reconstruct. He had also attended a symposium in America about MoM hips at which other members of the orthopaedic community provided evidence of similar experiences, and that influenced his thinking. He said that it was at the forefront of his mind that one had to be “*careful in getting them early, if there was a real problem*”. This was not an atypical attitude, and it is entirely understandable.
742. The metal ions test result, when it arrived on 30 September 2013, showed that the levels of chromium and cobalt in the blood remained very much lower than the levels that might be regarded as cause for concern— only 1.03 ppb and 0.84 ppb respectively. These measurements are even lower than Mr Nargol’s very low threshold of 2ppb for a normally functioning prosthesis.
743. The MRI scan also pointed away from a diagnosis of ARMD. The consultant radiologist said there was “*no significant joint effusion, fluid within the iliopsoas tendon sheath or evidence of soft tissue mass around joint with low signal intensity debris*”. On the contrary, he pointed to left abductor tendinosis at its trochanteric insertion, with calcification/ossification within the tendon, and problems with the right hip. Dr Ostlere and Dr Wilson agree that the images provide no evidence of a soft tissue mass or of any significant joint effusion. Dr Wilson was emphatic in his view that there was nothing in the imaging to support significant adverse soft tissue reaction to metallic debris, and he agreed with the view of the original radiologist about abductor tendinosis.
744. Dr Ostlere accepted that the aspects of the images relied upon by Dr Wilson in that regard were consistent with swelling or oedema but he thought that metal artefact had distorted the image. Dr Wilson’s response in cross-examination was compelling – he had carried out checks for distortion and was satisfied, for the reasons he explained, that there was none. The images – even those relied on by Dr Ostlere in re-examination - consistently supported Dr Wilson’s position and I prefer his evidence on this point to that of Dr Ostlere. So far as the orthopaedic experts were concerned, Mr Kim’s opinion was that the imaging pointed away from a diagnosis of ARMD. Professor Kay said the imaging “*had not revealed anything of significance*”.
745. The claimants sought to rely on the presence of some lucency or erosion around the stem and a small finger-like projection arising from the proximal erosion extending into the greater trochanter, which had been shown in X-rays but not the MRI scans. Mr Kim’s evidence was that the lucency next to the stem could be due to particulate debris, but it could also be due to mechanical causes if the stem had been inserted at an inward angulation. Even if it was caused by a reaction to metal, it was not a factor that would have prompted revision (and it was not, in fact, a factor on which Mr Herlekar had relied). Dr Wilson said, and I accept, that the radiological evidence was very equivocal as to whether the images showed granulation tissue due to mechanical loosening or an adverse reaction to ARMD. However, the images that were taken 12 months apart showed no progression, which pointed away from ARMD.

746. Mr Herlekar agreed that tendonosis could cause pain and discomfort, and that it was not an unusual finding for someone who had had the Hardinge approach, or for a sportsman such as Mr Woods. However, he said the pain would be a lateral pain rather than a pain felt in the groin. Whilst he agreed that both the blood tests and the MRI scans pointed away from a finding of ARMD, he said that neither test was 100% certain. He said it was 80% certain in the case of the blood tests and 85%-90% in the case of the MRI scan.
747. So far as the aspirate was concerned, around 3ml of fluid was taken from around the hip, which was consistent with the findings on the MRI imaging that there was no sign of any effusion or collection around the hip. There was no record of any technical difficulty being encountered in carrying out the aspiration. In one of the less satisfactory passages in his oral evidence, in which he appeared to be working backwards from an assumption that Mr Woods did have ARMD and finding features to justify that assumption, Professor Kay sought to suggest that because the 3mls of fluid from the hip joint was “*easily*” aspirated, that suggested there may have been more fluid in the joint to aspirate. This was clutching at straws. Mr Herlekar, who carried out that procedure, noted that the aspiration “*showed only a small amount of fluid*” which would be an odd remark to have made if he thought there was more fluid present. Given that Professor Kay agreed elsewhere in his evidence that larger amounts of aspirate were no cause for concern, his evidence about the 3ml of fluid extracted was a little difficult to follow. Ultimately, he said 3ml was probably slightly on the higher side of average, but he would not set great store by it.
748. A sample of the hip fluid was tested for metal ions, in the same way as Mr Nargol did in the case of Mrs Stalker. When the final decision to proceed to revision was made, Mr Herlekar described the cobalt and chromium levels from the joint fluid as “*significantly high*”. Those results appear to have played a significant part in Mr Herlekar’s decision to recommend revision. He described the levels of metal ions in the fluid as high, despite the fact that there was no scientific yardstick by which to measure them as high, medium or low. Mr Herlekar readily accepted, when it was put to him in cross-examination, that it was an error on his part to take those levels into account, given that (a) one would expect to find metal ions in the fluid around a metal hip (b) there was and is no scientific paper establishing that there is any connection between the level of metal ions in such fluid and ARMD.
749. Mr Herlekar also accepted that he had applied a lower threshold for revision than the MHRA guidelines. Mr Woods’ recollection was that Mr Herlekar was worried that he was having an adverse reaction to metal, and had advised him that his hip was “*showing the early signs of failure*”. Even though Mr Woods had a long history of fluctuating pain in his groin, and it had previously been suspected that he was suffering from referred pain from his lower back, there was no discussion between them of the possibility that something else might be causing the pain in his hip.
750. Moreover, Mr Woods had been doing some researches into MoM hips on the internet following his recall, and in consequence he became very worried about the fact that he had such a prosthesis. He was anxious to have something done about it. Therefore, this was a classic example of a combination of a patient becoming anxious to undergo a revision operation in consequence of the sensationalist media reporting about MoM hips, and a surgeon who was over-concerned about ARMD and would revise out of an abundance of caution despite the absence of any objective justification for doing so.

751. When he carried out the revision, Mr Herlekar noticed some metal staining but “*no huge ALVAL reaction*”. Mr Herlekar clarified that he was using “ALVAL” as a synonym for “ARMD” and that what he meant by this was that “*there was very early stage sort of blackening of the tissues*”, i.e. metallosis. In his oral evidence he confirmed that he found no tissue necrosis. He was surprised, because he was expecting to find evidence of ARMD and apart from the tissue staining there were no signs of it. He agreed that the 12 to 20 mls of brown fluid he recorded finding during surgery was not a surprising finding, and that it could have been caused by a haemorrhage, which was the view of Mr Kim. The tissue samples sent for histological examination were described as “*heavily blood stained*”. I am satisfied that the colour of the fluid was not an indication of ARMD.
752. There is a reference to “*necrotic debris*” in the contemporaneous histopathology report. It is unclear what that expression is referring to; Professor Nelson came up with a plausible theory that this was a reference to something that turned out to be fibrin when the slides were subject to a higher degree of magnification. On the basis of his images, I consider that explanation is probably correct.
753. The histology samples revealed no soft tissue damage. The histopathologist reported that they could see “*no evidence of the peripheral soft tissue or perivascular lymphocytic infiltrate expected in ALVAL*.” Nothing can be deduced from the presence of some macrophages, which was consistent with Mr Woods’ underlying osteoarthritis. The macrophage response was not an exuberant one, and not many metal particles were visible within them. As Professor Nelson explained, it was noteworthy that the macrophages and multinucleated giant cells visible on the cells were surrounded by many viable cells, which would not be the case if the tissue was dying. Moreover, there was no evidence of significant inflammation in the deep tissue. As Professor Athanasou confirmed, the presence of metallosis or visible metal particles tells one nothing about whether there has been a reaction to the metal, let alone an adverse one.
754. The only evidence that there was necrotic tissue in the samples taken comes from Professor Freemont, who found a “*widespread necrosis*” that neither the original histopathologist, Dr Theaker, nor Professor Nelson identified. I have already given my reasons for finding Professor Freemont’s evidence unreliable. I prefer the evidence of Professor Nelson, who carefully reviewed the slides at high magnification and confirmed that no appreciable tissue necrosis was present. His findings were consistent with the contemporaneous report. Mr Woods’ case was the lead claim in which Professor Freemont went so far as to suggest that Professor Nelson had been “*highly selective*” in his high magnification images. I have no reason to suppose that Professor Nelson had been anything of the kind. The accusation was a knee-jerk reaction by Professor Freemont to being proved wrong. An image that he said showed tissue necrosis had been subjected to high magnification, and he was forced to agree in cross-examination that the enlarged image showed viable blood vessels and viable macrophages and other cells in the tissue, as well as fibrin.
755. Following the revision surgery, Mr Woods made steady progress. At his three-week check-up with his GP he was noted to be mobilising well with one elbow crutch. When he next saw Mr Herlekar, on 24 March 2014, the latter said he had recovered “*remarkably well*”. The X-rays were satisfactory, and he had satisfactory left hip movement, though there was some discomfort in his left knee. By the time of his next



consultation on 19 May 2014, he could perform an active straight leg raise and full abduction without any problems. There was nothing of significance at his next consultation with Mr Herlekar in early November, but about three weeks after that, he saw his GP who recorded that he was “still having pain symptoms in the left hip” described as radiating down his left knee from the buttock.

756. In consequence of this pain, Mr Woods’ next appointment with Mr Herlekar was brought forward to 7 April 2015. Mr Woods told him that the pain on the left had significantly eased, but he continued to get some pain in the hip that he described as “*radiating to the back of his thigh as well as to the back of the knee joint*”. After reviewing the X-rays, Mr Herlekar concluded that most of his symptoms were originating in the back and that if they flared up he would need investigating for back problems. A year later, Mr Woods reported to Mr Herlekar that he still got some dull ache pain, but overall, he was happy with the outcome. He was discharged from the clinic.
757. When Mr Woods was examined by Mr Kim in October 2016 he was still complaining of intermittent symptoms, localised to the groin and the lateral aspect of his hip. His symptoms were aggravated by activity but were unpredictable.
758. Since signing his witness statement on 8 December 2016 Mr Woods has seen his GP for pain in his right hip. He was referred to a consultant, though he had not seen anyone by the time he came to give his oral evidence. The entry in the GP records states that this pain was similar to the left-sided pain before revision, and Mr Woods confirmed this in his oral evidence. The GP thought the location of the pain was possibly more indicative of a back problem.
759. At trial Mr Woods said that for the previous five or six months, his left hip was “*the best it has been for years*” and that he had “*a few twinges now and then*”. Contrary to the impression portrayed in his witness statement, the revision has had no appreciable impact on his life or his activities.
760. The evidence in this case falls a long way short of establishing that Mr Woods suffered from ARMD. On the contrary it seems highly likely that his MoM prosthesis was functioning perfectly well and would have continued to do so without problems had it been left alone. The blood metal levels were so low that one would not expect an adverse immunological reaction, unless he was one of those very unfortunate individuals who experienced extreme hypersensitivity – in which case, one would expect to see signs of adaptive immunological reactions. Professor Kay agreed that his metal ion levels “*no way suggested that there was a risk towards ARMD.*” There was no ALVAL type reaction here and the histopathology and radiology are inconsistent with ARMD. I accept Professor Nelson’s conclusion that Mr Woods’ histopathological findings are unremarkable in the context of an arthritic joint capsule with a MoM prosthesis.
761. It is unnecessary to make a specific finding as to the cause of Mr Woods’ symptoms of pain and discomfort in the hip, though Mr Kim made out a very persuasive case for tendinopathy as one of the causes. This fitted with Dr Wilson’s view that the MRI scan revealed some swelling/oedema by the adductors which was caused by tendinosis. Although Mr Kim and Professor Kay were in agreement that there was no evidence of active tendinosis, Mr Kim did explain that the pathological process was

one of impaired healing and therefore the problem tends to be a grumbling long-term chronic problem rather than an acute problem.

762. For all the above reasons, I find that Mr Woods did not suffer from ARMD. If I had reached the conclusion that he *did* suffer from ARMD, he made a remarkably good recovery from the revision operation, and the pain he has experienced since then is more likely than not to relate to his co-morbidities.