

# CeraNews

TO THE  
POINT

HEALTH ECONOMICS & POLICY

OUTCOMES RESEARCH

IMPLANT MATERIAL



## Can We Afford NOT to Use New Technologies?

By Stefan W. Kreuzer MD, MSc

## Medical Decision Making – Where Intuition and Evidence Meet

By Dr. Niklas Keller

## Identifying the Best Evidence for Ceramic Performance in THA

By Dirk Stengel MD, PhD, MSc

## Prevention Is Better than Cure: Challenges in the Treatment of PJI

By Prof. Rihard Trebše MD



### Video on Demand - Orthopedics Meets Dentistry 2022

What Drives Medical Intuition



### Supplement - Fact Box

Predicted Number of Preventable Complications and Surgical Revisions in THA  
Conditional on the Use of DoP and MoP



## Can We Afford NOT to Use New Technologies?



**Stefan Kreuzer MD, MSc**  
**Innov8 Orthopedics**  
**Houston, TX, US**

The United States' current spend on healthcare equates to 18% of GDP or USD 3.81 trillion and it is predicted rise to 19.7% by 2028 equating to USD 6.2 trillion. The trend is similarly reflected by most other countries and is simply unsustainable. Healthcare costs must be reduced in both the short- and long-term.

In the short-term, I envision a more optimized health delivery system, one where AI and machine learning enable physicians to incorporate the power of large data sets into personalized patient care. We know every patient is unique, and as is eloquently pointed out by Dr. Keller, the importance of intuition in medical decision-making should also not be underestimated. The synergy of this data driven personalized care will ultimately result in the most appropriate treatment plan within the most appropriate timescale.

Optimization also extends to reducing current resource burdens on the healthcare system, such as periprosthetic joint infection. Both the American and Australian registry annual reports 2022 highlight infection as the most common reason for revision in both hip and knee arthroplasties. The challenging diagnosis and treatment mean that any modifiable risk factors, such as metal bearings, should be avoided. Optimized healthcare delivery is a great value proposition which allows us to provide improved and more cost-effective care to more people in the immediate future.

When considering the long-term, we must be aware that the investment needed in technology is required today, not tomorrow. Robotics, navigation, and more durable devices will all assist in the reduction of future costs by reducing the revision burden for arthroplasty patients. The challenge, however, is predicting the long-term benefit in the short-term. Randomized Controlled Trials, as Prof. Stengel also recognizes, are not the be all and end all as ideal predictors for long-term benefits, despite being frequently required by state agencies or healthcare systems for acceptance.

Better methodologies to assess big data are growing quickly in our field and include the evolution of registry data using machine learning algorithms as well as predictive analytics and AI models which can be used to determine the potentials benefits of new technology. It stands to reason that a tool which is more precise, such as robotics or navigation and a bearing surface that reduces not only wear, but the risk of revision for infection and eliminates metal sensitivity should always be the default and not the exception. In this I can only agree with Prof. Trebše, prevention of infection is an absolute priority.

Much of the research in the field of human longevity and lifespan is based on optimization and personalization to improve health and extend life. David Sinclair, author of 'Lifespan'



hypothesizes that in the next ten years we will have the knowledge of how to extend life significantly. If these hypotheses become a reality a hip replacement would have to be durable enough to potentially last 50 years. While investment in new technology is costly, the current trends in spending are ultimately unsustainable going forward. The question should not be "Can we afford to use these new technologies?", the question should be "Can we afford NOT to use these new technologies?".

As a final note, if the patients were to be objectively presented with all the current evidence of new technologies, it is my belief that most of them would choose to receive a knee replacement performed with robotics, a hip replacement optimized with navigation, and consolidated bearing surfaces that have the potential to last a lifetime, such as ceramic-on-ceramic. So why the hesitant uptake of these technologies? I'll let you answer that question.

Stefan Kreuzer MD, MSc

# Medical Decision Making – Where Intuition and Evidence Meet

By Dr. Niklas Keller



**Dr. Niklas Keller**  
Simply Rational GmbH  
Berlin, Germany

We humans are not very good at making decisions. At least that is the classical view. We oversimplify complex matters, are constantly overwhelmed by too much information, and suffer from selective attention.<sup>1</sup> Our intuitive decisions are subject to systematic and predictable biases causing deviations from the rationally optimal choice.<sup>2,3</sup> These kinds of cognitive traps can influence how safety and medical risk are perceived and how medical decisions are made.

Everyone has biases. One much-studied example of a bias that has been shown significantly to impact medical decision-making is **base rate neglect**.<sup>4,5</sup> For example, when asked about the likelihood of a disease given a positive test result (positive predictive value), physicians often equate this with the sensitivity of the test and tend not to think of the prevalence of the disease as a factor. Consequently, the likelihood of a patient having a disease given a positive test result is often over-estimated.<sup>6</sup> This effect can be observed especially in the context of screening, since in screening, a population with very low disease prevalence is tested. How can evidence-based decision making be possible when human minds can find it difficult to convert this evidence into better decisions?

The classical view of human decision making is largely based on the 'Heuristics & Biases' research program founded by Daniel Kahneman and Amos Tversky in the late nineteen sixties.<sup>7</sup> This program locates the difficulties of sub-optimal judgement and decision making largely in the minds of decision makers and, critically, considers biases to be hardwired in human brains.<sup>8</sup> Subsequently, Kahneman thought that humans would not be able to 'de-bias' themselves: Like visual illusions, the knowledge of a bias does not protect one from falling into its trap.

## Base Rate Neglect

*Forgetting the underlying frequency (base rate) of an event, e.g., the prevalence of a disease, when making decisions in the light of new individuating evidence, like the results of a medical test.*<sup>4,5</sup>

## Decision making is not one size fits all

In the last thirty years a more nuanced and optimistic picture has emerged: decision making is best described as a pair of scissors where one blade is the structure of the task environment (the decision-ecology) and the other blade the strategy of the decision maker.<sup>9</sup> Only when the two are aligned can the scissor 'cut' the problem effectively. There is no 'universally rational' approach to decision-making. Instead, any strategy is only as good as its adaptation to the



specific decisional context, *i.e.*, its **ecological rationality**.<sup>10,11</sup> Consequently, one can improve decision making by either increasing the repertoire of strategies or competencies available to decision makers or by changing the decision environment. For example, the reason why the Romans could not multiply or divide in their heads was not because they had 'multiplication bias', but because the Roman numeral system (*i.e.*, the information environment) does not lend itself to such operations.

The same applies to base rate neglect: quite contrary to being hard-wired, as proponents of the Heuristics & Biases approach, base rate neglect practically disappears when critical information (sensitivity, specificity, and prevalence) is presented in the form of natural frequency trees<sup>6</sup> or if people are taught to convert conditional probabilities (such as sensitivity and specificity) into natural frequencies in their heads.<sup>12</sup> Such skills are now taught at medical schools in Germany (*e.g.*, Charité) and across the world (*e.g.*, Oxford) and have improved probabilistic test interpretation as a consequence.<sup>6,13</sup>

This research has also shown that **heuristics**, the cognitive shortcuts, or simple rules of thumb that people use to make decisions, do not only or even mostly lead to biases and sub-optimal decisions. Instead, it was demonstrated in many real-world scenarios outside of psychologists' laboratories, that intuitive heuristics can outperform far more complex and information-intensive strategies, even the latest AI technologies.<sup>14,15,16,17</sup> The discovery of these so-called "less-is-more-effects" (that one can make better decisions with less information and complexity) is seen as one of the most important findings in decision science in the last 30 years.<sup>18</sup>

### How to make a good decision

Gerd Gigerenzer, former Director of the Max-Planck Institute for Human Development in Berlin is the main proponent of **the 'Smart Heuristics' approach**. The key finding is: if there is a good understanding of the problem and a stable environment, then a problem can be quantified well, and more data and more complexity will lead to better outcomes.

In this situation, simplifying heuristics will always be second-best. However, if

there is a situation in which there is an incomplete understanding of the problem, or there is an unstable, dynamic environment, or good information is simply not available, then simplicity can lead to better outcomes. Here, the simple heuristics that underlie human intuitions can outperform AI because they focus on only the most relevant pieces of information and ignore the rest. This makes them less susceptible to the uncertainties and fluctuations of an uncertain world.<sup>20</sup>

#### The 'Smart Heuristics' approach

*"If risks are known, good decisions also require intuition and smart rules of thumb"<sup>19</sup>*

In an age of Evidence-Based Medicine (EBM), medical professionals are caught somewhere in the middle. On the one hand, new evidence is generated at an increasing rate and for many situations, high-quality data exists to inform medical decision-making. **The data on complication rates for different material components in THA presented in this issue are an example: taking the total number of revisions across all side-effects and patient types, data suggest that ceramic components offer better outcomes than their metal counterparts.** This is what Gigerenzer would refer to as "risks that are known". The important thing with known risks is to communicate them in ways that are transparent and understandable to decision makers, including both experts and laypersons.<sup>21</sup> The fact box format shown in the figure is such a way of transparently presenting the benefits and harms of medical treatments and represents current best-practice in risk communication.<sup>22,23</sup>

On the other hand, despite the available evidence, medical decision-making continues to be beset by a myriad of "risks that are unknown". How far is the scientific evidence applicable to the individual patient? What about new technologies and pharmaceutical interventions for which no or comparatively little evidence yet exists? In such situations, clinical intuition will continue to be a necessary and important part of medical decision-making. Necessary, because there simply is no data available for these questions, and important, because the research has shown that under these conditions, human expert intuition in fact does an excellent job and will not be easily replaced.

### The (un)reliability of intuition?

It is also important to note however, that human intuition is only as good as the information environment, in which it has developed. If certain types of information and feedback about the outcomes of a decision-maker's actions systematically do not reach them, then their intuition will be skewed. The asymmetric feedback about false-positives/over-diagnosis vs. false negatives/under-diagnosis is one of the factors that influences many medical professionals' intuition to "better treat too much than too little", termed **intervention-bias**.<sup>24</sup> Note that, contrary to the traditional view, this bias is not hardwired into human minds but rather the result of an adaptive mind attuned to a biased information environment. A similar effect can be observed if side-effects of treatments systematically differ with respect to how long they take to develop and whether they surface within the same specialty as the one responsible for the original intervention.



Fact Box

Sources: NJR – Bespoke Implant Report for Ceramtec - Biolox Delta articulations versus metal on polyethylene 25-10-2022

## Predicted number of preventable complications and surgical revisions in THA conditional on the use of BIOLOX®delta-on-Polyethylene (DoP) compared to Metal-on-Polyethylene (MoP)

The table shows the incidence of revision for various complications with metal femoral heads on polyethylene liners (MoP) per 10,000 primary Total Hip Arthroplasties (THA) observed in the National Joint Registry (NJR) of England, Wales, Northern Ireland, the Isle of Man and Guernsey. Using NJR data, expected and observed differences, and a multivariate regression model (controlling for age, gender, American Society of Anaesthesiologists (ASA) physical status classification, year of implantation, stem and cup fixation, and head size) it was estimated how many complications and revisions could have been prevented if all patients with MoP had received BIOLOX®delta femoral heads on polyethylene liners (DoP).

Note the estimates come from a statistical model not allowing for causal inference.

CAUSE OF REVISION*	Incidence of revision per 10,000 primary THAs with MoP bearings	Estimated incidence of revision per 10,000 primary THAs given a DoP had been used rather than MoP
<b>Infection</b>	<b>55</b>	<b>42</b>
<hr/>		
<b>Periprosthetic Fracture</b>		
- Stem	65	53
- Socket	6	5
<hr/>		
<b>Dislocation</b>	<b>73</b>	<b>64</b>
<hr/>		
<b>Aseptic Loosening</b>		
- Stem	39	26
- Socket	52	28
<hr/>		
<b>Unexplained Pain</b>	<b>22</b>	<b>17</b>
<hr/>		
<b>TOTAL ACROSS ALL COMPLICATIONS</b>	<b>302</b>	<b>231</b>

\*Multiple complications may lead to revision.

### What does this mean in summary?

In summary, the data from NJR appear to show that 71 per 10,000 revisions for major complications may be prevented by opting for BIOLOX®delta-on-Polyethylene rather than Metal-on-Polyethylene bearing surfaces in primary Total Hip Arthroplasty. Estimates were controlled for key independent demographic and therapeutic confounding variables.

Last update: November 2022

<https://www.ceramtec-medical.com/en/infocenter>

The data used for this analysis were obtained from the National Joint Registry ("NJR"), part of the Healthcare Quality Improvement Partnership ("HQIP"). HQIP, the NJR and/or its contractor, NEC Software Solutions (UK) Limited ("NEC") take no responsibility (except as prohibited by law) for the accuracy, currency, reliability and correctness of any data used or referred to in this report, nor for the accuracy, currency, reliability and correctness of links or references to other information sources and disclaims all warranties in relation to such data, links and references to the maximum extent permitted by legislation including any duty of care to third party readers of the data analysis. The summary implant reports are available upon request: [a.porporati@ceramtec.de](mailto:a.porporati@ceramtec.de).

### The 'hot-stove' effect: The perceived unsafety

Recent research further indicates that human beings react very differently to information presented in written format ('Decisions from Description' such as a fact box) than to information they have personally experienced ('Decisions from Experience', such as complications directly experienced or even anecdotal evidence from colleague<sup>25</sup>). If human beings experience a rare event, they are likely to heavily overweigh its likelihood, no matter the descriptive evidence presented. This is known as a **'hot-stove'** effect, *i.e.*, the avoidance of actions for which one has experienced negative outcomes even only once in the past. It is very difficult to overcome such effects with mere descriptive information. However, formats have been developed, which allow professionals to experience the frequency of events in simulated environments. Such interventions can potentially over-come hot-stove or other experience-based effects and thus positively impact medical decision-making in the direction of the best available evidence.<sup>26</sup> A smart communication of the available evidence can therefore not just be told to people - in some cases, it must be experienced, either through exchanges with colleagues or in simulated environments which reflect the true underlying statistical frequencies of events.

This is known as a **'hot-stove'** effect, *i.e.*, the avoidance of actions for which one has experienced negative outcomes even only once in the past. It is very difficult to overcome such effects with mere descriptive information.

### Integrating medical evidence and clinical intuition

Even in today's world of Evidence-Based Medicine, good decision making will continue to require both medical evidence and the clinical intuition of experienced physicians. Generally, it is difficult for humans to switch off their intuition completely. This makes it ever more important that possible biases and asymmetries inherent in learning and information environments are critically reflected. Otherwise, those biases will be mirrored by the human mind, often unnoticed, and this may negatively impact clinical intuitions and the decisions based on them.





1. Kahneman D. *Thinking, Fast and Slow*. Farrar, Straus and Giroux Publishing; 2011
2. Wilke A, Mata R. Cognitive bias. In: Ramachandran VS, ed. *The Encyclopedia of Human Behavior, Vol. 1*. Academic Press; 2012:531-535.
3. Gilovich T, Griffin D, Kahneman D. *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press; 2002.
4. Bar-Hillel M. The base-rate fallacy in probability judgments. *Acta Psychol*. 1980;44(3):211-233. doi:10.1016/0001-6918(80)90046-3.
5. Welsh M, Navarro D. Seeing is believing: Priors, trust, and base rate neglect. *Organ Behav Hum Decis Process*. 2012;119(1):1-14. doi:10.1016/j.obhdp.2012.04.001
6. Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping doctors and patients make sense of health statistics. *Psychol Sci Public Interest*. 2007;8(2):53-96. doi:10.1111/j.1539-6053.2008.00033.x.
7. Kahneman D, Tversky A. On the psychology of prediction. *Psychol Rev*. 1973;80(4):237-251. doi:10.1037/h0034747.
8. Ariely D. *Predictably Irrational: The hidden Forces that Shape our Decisions*. HarperCollins Publishers; 2008.
9. Simon AH. Invariants of human behaviour. *Ann Rev Psychol*. 1990;41:1-19. doi:10.1146/annurev.ps.41.020190.000245.
10. Goldstein DG, Gigerenzer G. Models of ecological rationality: The recognition heuristic. *Psychol Rev*. 2002;109(1):75-90. doi:10.1037/0033-295x.109.1.75.
11. Todd PM, Gigerenzer G. *Ecological Rationality: Intelligence in the World*. Oxford University Press; 2012. doi:10.1093/acprof:oso/9780195315448.001.0001.
12. Gigerenzer G, Hoffrage U. How to improve Bayesian reasoning without instruction: Frequency formats. *Psychol Rev*. 1995;102(4):684-704. doi:10.1037/0033-295X.102.4.684.
13. Jenny MA, Keller N, Gigerenzer G. Assessing minimal medical statistical literacy using the Quick Risk Test: a prospective observational study in Germany. *BMJ Open*. 2018;8(8):e020847. doi:10.1136/bmjopen-2017-020847.
14. DeMiguel V, Garlappi L, Uppal R. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev Financ*. 2009;22:1915-1953. doi:10.1093/rfs/hhm075.
15. Wübben M, v. Wangenheim F. Instant customer base analysis: Managerial heuristics often get it right. *J Mark*. 2008;72:82-93.
16. Gigerenzer G, Goldstein DG. Reasoning the fast and frugal way: Models of bounded rationality. *Psychol Rev*. 1996;103(4):650-669. doi:10.1037/0033-295x.103.4.650.
17. Gigerenzer G, Gaissmaier W. Heuristic decision making. *Annu Rev Psychol*. 2011;62:451-482. doi:10.1146/annurev-psych-120709-145346.
18. Brighton H, Gigerenzer G. Homo heuristicus: Less-is-more effects in adaptive cognition. *Malays J Med Sci*. 2012;19(4):6-16.
19. Gigerenzer G. *Risk savvy: How to make good decisions*. Viking Publishers; 2014.
20. Gigerenzer G, Brighton H. Homo heuristicus: Why biased minds make better inferences. *Top Cogn Sci*. 2009;1(1):107-143. doi:10.1111/j.1756-8765.2008.01006.x.
21. Gigerenzer G, Muir-Gray JA. *Better doctors, better patients, better decisions: Envisioning health care 2020*. The MIT Press; 2011.



22. Brick C, McDowell M, Freeman ALJ. Risk communication in tables versus text: A registered report randomized trial on 'fact boxes'. *R Soc Open Sci*. 2020;7(3):190876. doi:10.1098/rsos.190876.
23. McDowell M, Gigerenzer G, Wegwarth O, Rebitschek FG. Effect of tabular and icon fact box formats on comprehension of benefits and harms of prostate cancer screening: A randomized trial. *Med Decis Making*. 2019;39(1):41-56. doi:10.1177/0272989X18818166.
24. Foy AJ, Filippone EJ. The case for intervention bias in the practice of medicine. *Yale J Biol Med*. 2013;86(2):271-280.
25. Hertwig R, Wulff DU. A description-experience framework of the psychology of risk. *Perspect Psychol Sci*. 2022;17(3):631-651. doi:10.1177/17456916211026896.
26. Wegwarth O, Ludwig WD, Spies C, Schulte E, Hertwig R. The role of simulated-experience and descriptive formats on perceiving risks of strong opioids: A randomized controlled trial with chronic noncancer pain patients. *Patient Educ Couns*. 2022;105(6):1571-1580. doi:10.1016/j.pec.2021.10.002.

# Best Evidence Synthesis of the Clinical Performance of Ceramic Bearings in Total Hip Arthroplasty with Focus on BIOLOX<sup>®</sup>delta

By Dirk Stengel MD, PhD, MSc



**Dirk Stengel MD, PhD, MSc**  
BG Kliniken - Klinikverbund  
der gesetzlichen  
Unfallversicherung gGmbH  
Berlin, Germany

A key principle and driver of global economy is that any company aims to produce a product far better or much cheaper (while of similar quality and function) than that of its competitors to foster its role as a market leader. Both the pharmaceutical and medical device industry have a unique societal role and responsibility in this global competition, as their products may immediately affect the fate of an individual patient, as well as health-related outcomes and function of a population with a certain disorder or injury.

Estimating the **utility and value** of a health-care intervention demands a thorough trade-off between its reported benefits and harms. Unbiased, transparent, and easy, understandable information will decide whether payers, professionals and patients opt for a certain treatment over another, even if it is associated with higher tangible (*i.e.*, monetary) and intangible costs (*e.g.*, a higher risk of adverse events).

Tribological pairing represents one among an infinite number of variables affecting outcomes after total hip arthroplasty (THA). Demographic baseline profiles, comorbidity, surgical expertise and (minimally-invasive) access routes, navigation, hardware from various manufacturers, cemented or cementless fixation, peri-operative and rehabilitation protocols, and many other factors may have a far greater impact on recovery, function, and long-term revision-free component survival than the individual material constituting the acetabular liner and femoral head.

Keep in mind that most novel treatments, especially in orthopedics, represent step innovations with marginal effect sizes. Showing a difference to the standard of care or other therapeutic options, controlling for multiple confounding items, needs thousands of subjects and datasets.

## What is the best available evidence?

Health-care authorities such as the German Institute for Quality and Efficiency in Health Care (*Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, IQWiG*) always pose the following key questions, all of which must be answered in detail to increase the likelihood of coverage by insurers and acceptance by providers:

1. How valid is the current best available evidence on the safety and effectiveness of a therapeutic intervention (here: ceramic bearings) compared to the standard of care to draw meaningful conclusions, ideally to draw causal inferences?



2. Is there scientific information from data sources with a low risk of bias, specifically well designed (!) randomized controlled trials (RCT) with a sample size large enough to make robust predictions?
3. Does the reported effectiveness or benefit in patient-centered outcomes (*e.g.*, function, health-related quality of life) outperform any intervention-specific risk (here: audible noise and/or squeaking, ceramic fractures)?
4. If there is any measurable difference between interventions, is it both (i.) statistically significant (*i.e.*, beyond the play of chance) and (ii.) clinically relevant (*i.e.*, above a certain threshold recognizable by patients)?

There is minimal consensus among scientists and health-care professionals that a potentially innovative, useful and valuable intervention requires, at least, **a biologically reasonable mode of action** (demonstrated by reproducible pre-clinical or animal experiments) which, in theory, may increase the likelihood of better long-term clinical outcomes compared to the standard of care or other thinkable therapeutic options.

Long story short- almost **no** novel or modern drug or medical device meets or is even close to meeting all criteria.

The buzzword **Evidence-**

**Based Medicine (EBM)** is used excessively but too often erroneously. The term was coined about 30 years ago by Canadian and British clinician scientists, depicting "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients".<sup>1</sup> A common misbelief of the still applicable basic principles of EBM is that its inventors and propagators equated **current best evidence** with RCT. Fundamental advancements in information technology, statistical methods, genomics and molecular biology, precision medicine, machine learning, open-access publishing, and many others, subsequently overcame the original concept, which nowadays must be considered historical and outdated.

This is specifically true for the well-known **evidence pyramid**. Bruce G. Charlton, a retired British medical doctor and professor of theoretical medicine was well known for his often controversial and even bizarre statements - but he was right in stressing "a hierarchy of methods is amazing nonsense".<sup>2</sup> This simply means that experimental or non-experimental set-ups must be adapted to the individual problem to be solved- there is no one-size-fits-all ranking of study designs. Also, a poorly planned, conducted and reported, small-sized RCT with unexplained post-randomization drop-outs etc. may be no better, solid, or meaningful than a case series, while results from a large-scale registry with nearly complete long-term follow-up may substantially influence clinical practice and health-care decisions.

More flexible approaches to assess the trustworthiness and relevance of scientific evidence are:

1. the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) instrument, continuously developed and evaluated by the international GRADE Working



Group (<https://www.gradeworkinggroup.org/>), which considers individual methodological features of a clinical investigation rather than its general design,

2. the second version of the Cochrane Risk of Bias Tool (RoB-2) for randomized trials (<https://methods.cochrane.org/risk-bias-2>).

Without doubt, **systematic reviews** (of individual studies or systematic reviews) remain the best source for informed decision making in health-care. If done properly, they may show the advantages and disadvantages of diagnostic tests and therapeutic interventions for a certain condition or disease in an unbiased fashion and bring them into the context of actual scientific and clinical standards. A chain is only as strong as its weakest link, and a systematic review can only be as good as its included individual studies and trials.

A **best evidence synthesis** may be characterized as a special form of a mixed-methods study. It combines principles of systematic reviews and meta-analyses, systematic reviews of systematic reviews, scoping (or narrative) reviews, and health-technology assessments. It usually covers a wide range of evidence (*e.g.*, systematic reviews and meta-analyses of RCTs and cohort studies, individual trials, registries, routine, and administrative data etc.), as derived from a formal, reproducible systematic search of the literature in multiple databases (*e.g.*, PubMed Medline, Ovid Medline and Embase, Cochrane Library, and grey literature), supplemented by a snowball procedure (*i.e.*, a search among related articles and cited references, prompting another search among related articles and cited references, until all related articles and cited references match).

### Best available evidence on BIOLOX®*delta*

CeramTec assigned an independent expert to compile the best available evidence on the effectiveness and safety of ceramic bearings focusing on BIOLOX®*delta* in THA. This included a reproducible search among different databases (*e.g.*, Ovid Medline and Embase), quality assessment, data extraction from original articles and aggregation using advanced statistical methods. Best evidence syntheses share features of so-called **living systematic reviews** which continuously incorporate and adapt objectives conditional on new information. Originally planned as a single comprehensive overview, multiple new questions arose after digging deeper and deeper into the available amount of data. Consequently, the work had to be split into three consecutive parts, each of which explored different data sources under specific scopes and perspectives.

Current experimental evidence from RCTs of ceramic-on-ceramic (CoC) and ceramic-on-polyethylene (CoP) compared to metal-on-polyethylene (MoP) turned out to be sparse and methodologically weak. To date, there is no large-scale, confirmatory multicenter RCT comparing CoC and / or CoP to alternative pairings in THA.

Unlike the U.S. Food and Drug Administration (FDA), health authorities in Europe, particularly in Germany, tend to be disobliging to working with industry in developing sound but feasible study designs to answer questions of importance to both manufacturers and the health care system. The IQWiG, Federal Institute for Drugs and Medical Devices (Bundesinstitut für Arzneimittel und Medizinprodukte, BfArM), and most Notified Bodies in Germany share the position that it is solely the industry's responsibility to deliver the necessary clinical evidence for decision making, without providing specific guidance how this should be done. Absence of evidence of a benefit given the lack of RCTs is not evidence of absence of a benefit by other data sources. Producers must, of course, prove their product is equally or more effective compared to the (leading) comparator on the market, while being safe. However, according to the famous philosopher Hans Albert, the so called first bridge principle reads "Shall implies Can". (Albert H. Traktat über kritische Vernunft. 5<sup>th</sup> Ed. Mohr, University of California, 1991.) This means health care authorities cannot insist on a proof of effectiveness by a large-scale RCT if a large-scale RCT is impossible to be conducted for comprehensible and transparently explained reasons.

However, current best available evidence from international joint registries gives strong indications and tendencies that BIOLOX<sup>®</sup>*delta* bearings (CoC and CoP) in THA are associated with a lower overall risk of revision, mainly driven by a lower risk of revision for periprosthetic joint infection.

Cumulative 2 to 13 year survival of THAs with BIOLOX<sup>®</sup>*delta* bearings range from 94 to 100%, accompanied by significant improvements in function and pain comparable to other couplings.

The incidence of audible noise and/or squeaking reported in clinical studies ranges from 1.6 to 6.8%, with heterogeneous definitions and assessment procedures. There is no consistent or statistically conspicuous association between noise and/or squeaking and pain, function, patient-reported outcome measures (*e.g.*, OHS or WOMAC), and revision rates.

The pathophysiology of audible noise and squeaking with CoC remains unclear. Of note, even if reported by patients, the phenomenon cannot always be reproduced during objective physical examination. There is no valid tool to measure or quantify noise in a longitudinal fashion, and, with the exemption of Australia<sup>3</sup>, noise is not routinely recorded in joint replacement registries worldwide. Noise does not signal ceramic fracture and is a rare cause of revision. One may even assume a so-called recall bias, as single case reports of CoC associated squeaking prompted patients after a long period without any complaints to report about noise sustained in the past.

Ceramic compounds, however, prevent biofilm formation, as substantiated by multiple laboratory experiments. The observed lower rate of revision for deep, implant-related infections thus has a pathophysiological explanation.



This best evidence synthesis suggests that:

- It is 15 to 33 times more likely ceramic bearings avoid a revision for infection than causing a revision for audible noise.
- It is 38 to 85 times more likely ceramic bearings avoid a revision for infection than causing a revision for ceramic head fractures.
- It is 3 to 6 times more likely ceramic bearings avoid a revision for infection than causing a revision for ceramic liner fractures.

This comprehensive review suggests a **favorable benefit-risk-ratio** of ceramic bearings (CoC and CoP) compared to other couplings in total hip arthroplasty (THA). While the lack of confirmatory evidence from large-scale randomized controlled trials (RCTs) cannot and must not be denied, registry data speak a clear language.

Ceramic components manufactured from alumina (BIOLOX<sup>®</sup>*forte*) and alumina matrix composite (BIOLOX<sup>®</sup>*delta*) have established themselves as durable bearings in total hip arthroplasty (THA), either as ceramic-on-polyethylene (CoP, with all its advancements like highly cross-linked polyethylene [HXLPE], and others), or ceramic-on-ceramic (CoC) couplings. Scientific literature puts emphasis on ceramic specific adverse events like squeaking (a phenomenon common to hard-on-hard bearings) and component fractures. However, ceramic bearings show a lower risk of revision for prosthetic joint infection, probably the most serious complication in total joint arthroplasty.



1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71-72. doi:10.1136/bmj.312.7023.71.
2. Charlton BG. Fundamental deficiencies in the megatrial methodology. *Curr Control Trials Cardiovasc Med*. 2001;2(1):2-7. doi:10.1186/cvm-2-1-002.
3. Owen DH, Russell NC, Smith PN, Walter WL. An estimation of the incidence of squeaking and revision surgery for squeaking in ceramic-on-ceramic total hip replacement: a meta-analysis and report from the Australian Orthopaedic Association National Joint Registry. *Bone Joint J*. 2014;96-B(2):181-187. doi:10.1302/0301-620X.96B2.32784.



# Prevention is Better than Cure: Challenges in the Diagnosis and Treatment of PJI and Metal as Modifiable Risk Factor

By Prof. Rihard Trebše MD



**Prof. Rihard Trebše MD**  
Valdoltra Orthopaedic  
Hospital Ankaran, Slovenia

Age-related degenerative joint diseases, such as osteoarthritis of the hip and knee, are a major cause of disability globally and a huge problem for healthcare systems worldwide.<sup>1</sup> Total joint arthroplasty is one of the most successful orthopedic procedures that has reduced pain and regained joint function in millions of osteoarthritis patients worldwide.<sup>2</sup> Nevertheless, 5-10% of patients undergoing total hip and knee arthroplasty will have to undergo revision surgery with the exchange of the prosthesis at 10 years after the primary procedure.<sup>3</sup>

## Aseptic or septic?

One of the most common reasons for revision of a joint replacement is aseptic loosening, which can be related to inadequate primary stability, mechanical weakening of fixation over time and particle-induced bone resorption.<sup>4</sup> Particles may be generated by wear, mechanically assisted corrosion at the taper junction, oxidation reactions and demineralization by pathogens. Metal ions released by biocorrosion products, although playing a subordinate role compared to wear particles, also increase the excretion of inflammatory markers, which can contribute to aseptic loosening. As millions of joint replacements are implanted each year and the patient population is increasingly active, the growth of revision surgeries is significant and represents a burden for the current healthcare system and even more so for the future. As such, it may not be sustainable henceforth, as revisions are associated with 3 to 8 times higher mortality rates, poorer clinical outcomes, and much higher costs than primary operations.<sup>3</sup>

One of the most serious complications leading to challenging revisions is extensive tissue damage caused by adverse local tissue reactions from metal debris (ARMD) and toxic metal ions.<sup>5</sup> Revision for ARMD can often end with a complication associated to major revision procedures, of which the most serious one remains the periprosthetic joint infection (PJI). Whereas PJI is cause of revision for 2-3% of primary procedures, it is also the reason for up to 10% of re-revision procedures.<sup>3</sup> It is disabling for patients, impairing quality of life, requires very invasive treatment, and there is a high risk of significant adverse events. The mortality rate of revisions for infection is five times that for aseptic reasons and the risk of sepsis in revisions is seven times higher compared to primary procedures, with consequent increase of morbidity and mortality.

One of the most serious complications leading to challenging revisions is extensive tissue damage caused by adverse local tissue reactions from metal debris (ARMD) and toxic metal ions.<sup>5</sup>



However, the number of infections may be significantly underestimated, as differentiation between aseptic and septic loosening remains difficult due to similar clinical symptoms and radiographic findings, so that a significant number of misdiagnoses are assumed.<sup>4,6</sup> Bacteria activate the host immune system and can cause implant loosening. Therefore, bone resorption with moderate pain and radiographic findings observed in aseptic loosening may also be the result of infection, usually caused by low-virulent bacteria. It follows that low-grade PJIs also referred to as subclinical or occult, may be unsuspected and with undetectable symptoms and as a consequence may contribute to prosthetic failure diagnosed as aseptic. Therefore, the diagnostic pitfall is a major current problem, and early detection of infection remains crucial to the successful treatment of PJI. Infection must always be ruled out in every case, as underdiagnosis risks adverse outcomes with inappropriate treatment aimed at aseptic revision when infection is present, often leading to subsequent invasive treatment. Unfortunately, correctly distinguishing infection from aseptic failure is a real problem in clinical practice and, more importantly, depends on the diagnostic threshold used, awareness and acute clinical suspicion.<sup>6</sup>

No single diagnostic approach has gained acceptance as a reference standard for clinical practice. Clearly, in this scenario, the best strategy may firstly be avoiding all possible modifiable risks in the primary procedure.<sup>6</sup>

### Diagnostic and therapeutic approaches of PJI

The predominant pathogens responsible for PJI are coagulase-negative staphylococci, staphylococcus aureus, streptococci, gram-negative bacilli, enterococci, and anaerobes.<sup>7</sup> In about 10% of PJIs, mixed infections are detected, and in an additional 10%, no microorganisms are detected by methods currently used.<sup>8</sup> This distribution has also been observed in Slovenian patients.<sup>9</sup>

Joint infection can present itself in many ways, ranging from fulminant joint sepsis with clear signs of infection to more indolent symptoms, such as pain or joint dysfunction. The mode of clinical presentation refers to the pathogenesis (planktonic bacteria vs. biofilms) and the microbial etiology of the infection (high vs. low virulence microorganisms). While fever and erythema are quite specific, their sensitivity for diagnosing of PJI is low.<sup>10</sup> Pain and reduced range of motion are the most sensitive clinical signs in PJI, but the specificity is low as they overlap greatly with aseptic failures. Most likely there is also an unknown percentage of PJI that does not manifest clinically but can become clinically important with time or if some other conditions influencing local or systemic immune status develop (*e.g.*, use of an immunosuppressive drug or similar).

As already stated, an accurate diagnosis is the starting point for effective treatment: underdiagnosis of PJI leads to inadequate treatment with serious consequences. Overdiagnosis results in inappropriate invasive treatment. Despite attempts to accurately define diagnostic

criteria for PJI, diagnostic and therapeutic approaches in the management of PJI remain challenging. High-grade infection is easier to recognize, while low-grade infection and so called 'aseptic' loosening are problematic.<sup>6</sup>

Due to many reasons including the complexity of PJI presentation, geographical variations of desired diagnostic accuracy, expensive testing availability, and disagreement about the accuracy of some of the included tests, a single definition for PJI has yet to be accepted. This would

Fig. 1:  
Three-level traffic light' EBJIS  
definition of PJI.<sup>13</sup>

	Infection unlikely	Infection likely	Infection confirmed	
	All findings negative 	Two positive findings 	Any positive finding 	
CLINICAL	Clinical features	<ul style="list-style-type: none"> <li>• Early radiographic loosening</li> <li>• Wound healing problems</li> <li>• Recent fever/bacteraemia</li> <li>• Purulence around prosthesis</li> <li>• CRP &gt; 10mg/l</li> </ul>	Sinus tract communication with the joint ± visualization of prosthesis	
	C-reactive protein			
LABORATORY	Synovial fluid	<ul style="list-style-type: none"> <li>• Leukocyte count ≤ 1500</li> <li>• PMN ≤ 65%</li> </ul>	<ul style="list-style-type: none"> <li>• Leukocyte count &gt; 1500</li> <li>• PMN &gt; 65%</li> </ul>	<ul style="list-style-type: none"> <li>• Leukocyte count &gt; 3000</li> <li>• PMN &gt; 80%</li> <li>• Positive Alpha-defensin</li> </ul>
	Microbiology	<ul style="list-style-type: none"> <li>• All cultures negative</li> <li>• No growth on sonication</li> </ul>	<ul style="list-style-type: none"> <li>• Single positive culture (aspiration or intra-operative)</li> <li>• &gt; 1 CFU/ml any organism on sonication</li> </ul>	<ul style="list-style-type: none"> <li>• ≥ 2 positive samples with the same microorganism</li> <li>• &gt; 50 CFU/ml any organism on sonication</li> </ul>
	Histology	Negative	Presence of ≥ 5 neutrophils in a single HPF	<ul style="list-style-type: none"> <li>• Presence of ≥ 5 neutrophils in ≥ 5 HPF</li> <li>• Visible microorganisms</li> </ul>
RADIOLOGY	Nuclear imaging	Negative 3-phase isotope bone scan	Positive white blood cell labelled scintigraphy	

©Rihard Trebše



provide a unique practical guide for clinicians in the PJI management, and it would allow comparison between the studies.

To overcome this problem, the European Bone and Joint Infection Society (EBJIS) has recently proposed the novel 'traffic light' approach (Figure 1), which divides patients by the likelihood of infection (green or amber) or confirmed infection (red).<sup>11</sup> The approach is not binary (*i.e.*, 'infection' or 'no infection') and it has shown clinically to detect a higher percentage of low-grade infections.<sup>8,12</sup>

### Metal as risk factor for PJI and differential influence of bearing wear debris on local tissue immunocompetence

Several risk factors for PJI following arthroplasty surgery have been identified on the patient side, such as male gender, elevated BMI, higher ASA grade and specific comorbidities. Consequently, with targeted interventions applied pre-operatively patients can be medically optimized in order to reduce the risk of PJI. On the implant side metal bearings are among the most important and easily modifiable risk factors associated with PJI and have been associated with a higher risk of revision for PJI.

Observational studies have shown that the bearing material has an influence on the incidence of revision for PJI: ceramic-on-ceramic bearings having the lowest and metal-on-metal bearings having the highest whilst metal-on-polyethylene and ceramic-on-polyethylene have an intermediate incidence of revision for PJI.<sup>14,15</sup> Epidemiological studies have found an association between PJI and the bearing material, but the causative mechanism has still to be determined. The pathogenic influence of metal-on-metal bearings has been widely investigated after THA patients developed adverse reactions to metal debris. In contrast ceramic-on-ceramic bearings are known for their high wear resistance and their excellent biocompatibility. Due to the very low concentration of ceramic wear particles produced *in-vivo* and their chemical-physical characteristics, an inappropriate inflammatory response in patients is very unlikely.

It may be speculated, that the incidence of clinically manifested PJI (especially low-grade PJI) may be influenced by the bearing surface type because of a differential influence of various bearing materials on local and systemic innate host defense response. In fact, when particulate debris of different bearing materials were added to human macrophages in tissue culture, it was found that macrophages reacted differently depending on the chemical composition and size of the particles. The ceramic composite particles were shown to induce no cytotoxic reaction in human macrophages, indicating the biological safety of the ceramic composite particles.<sup>4</sup> On the other hand, metallic particles resulted cytotoxic for human macrophages. In addition, metallic particulate has also shown to increase excretion of some pro-inflammatory cytokines, suggesting a high inflammatory potential of metal wear particles. We may therefore assume that metallic bearing materials could locally immunocompromise periprosthetic tissue and promote the onset of PJI that would have remained silent in a less toxic environment.



It can be concluded that a possible explanation for the higher risk of PJI with metal bearings may be related to the generated metal particles and their biocorrosion products, which could locally immunocompromise periprosthetic tissue and promote the flare of infection. Research on this hypothesis is ongoing.

### Prevention as best strategy against PJI

PJI is a real threat in THA, which can severely compromise patient health. Metal bearings have been identified as risk factor for PJI. Diagnosis and treatment are still very challenging representing a resource burden on the healthcare system, and any potential modifiable risk factor should be avoided a priori.



1. Briggs AM, Cross MJ, Hoy DG, et al. Musculoskeletal health conditions represent a global threat to healthy aging: A report for the 2015 World Health Organization world report on ageing and health. *Gerontologist*. 2016;56 Suppl 2:S243-S255. doi:10.1093/geront/gnw002.
2. Kamaruzaman H, Kinghorn P, Oppong R. Cost-effectiveness of surgical interventions for the management of osteoarthritis: a systematic review of the literature. *BMC Musculoskelet Disord*. 2017;18(1):183. doi:10.1186/s12891-017-1540-2.
3. Weber M, Renkawitz T, Voellner F, Craiovan B, Greimel F, Worlicek M, Grifka J, Benditz A. Revision surgery in total joint replacement is cost-intensive. *Biomed Res Int*. 2018:8987104. doi:10.1155/2018/8987104.
4. Hodges NA, Sussman EM, Stegemann JP. Aseptic and septic prosthetic joint loosening: Impact of biomaterial wear on immune cell function, inflammation, and infection. *Biomaterials*. 2021;278:121127. doi:10.1016/j.biomaterials.2021.121127.
5. de Steiger RN, Hatton A, Peng Y, Graves S. What is the risk of THA revision for ARMD in patients with non-metal-on-metal bearings? A study from the Australian National Joint Replacement Registry. *Clin Orthop Relat Res*. 2020;478(6):1244-1253. doi:10.1097/CORR.0000000000001277.
6. Villa JM, Pannu TS, Piuze N, Riesgo AM, Higuera CA. Evolution of diagnostic definitions for periprosthetic joint infection in total hip and knee arthroplasty. *J Arthroplasty*. 2020;35(3S):S9-S13. doi:10.1016/j.arth.2019.10.032.
7. Papalini C, Pucci G, Cenci G, Mencacci A, Francisci D, Caraffa A, Antinolfi P, Pasticcini MB. Prosthetic joint infection diagnosis applying the three-level European Bone and Joint Infection Society (EBJIS) approach. *Eur J Clin Microbiol Infect Dis*. 2022;41(5):771-778. doi:10.1007/s10096-022-04410-x.
8. Tande AJ, Patel R. Prosthetic joint infection. *Clin Microbiol Rev*. 2014;27(2):302-345. doi:10.1128/CMR.00111-13.
9. Rak M, Barlič-Maganja D, Kavčič M, Trebše R, Čör A. Comparison of molecular and culture method in diagnosis of prosthetic joint infection. *FEMS Microbiol Lett*. 2013;343(1):42-48. doi:10.1111/1574-6968.12125.
10. Shohat N, Goswami K, Tan TL, et al. Fever and erythema are specific findings in detecting infection following total knee arthroplasty. *J Bone Jt Infect*. 2019;4(2):92-98. doi:10.7150/jbji.30088.
11. McNally M, Sousa R, Wouthuyzen-Bakker M, et al. Infographic: The EBJIS definition of periprosthetic joint infection. *Bone Joint J*. 2021;103-B(1):16-17. doi:10.1302/0301-620X.103B1.BJJ-2020-2417.
12. Trebse R, Trampuz A. The definition of prosthetic joint infections (PJI). In: Trebse R, ed. *Infected total joint arthroplasty*. London: Springer; 2012:23-30.
13. McNally M, Sousa R, Wouthuyzen-et al. Infographic: The EBJIS definition of periprosthetic joint infection. *Bone Joint J*. 2021;103-B(1):16-17. doi:10.1302/0301-620X.103B1.BJJ-2020-2417.
14. Pitto RP, Sedel L. Periprosthetic joint infection in hip arthroplasty: is there an association between infection and bearing surface type? *Clin Orthop Relat Res*. 2016;474(10):2213-2218. doi:10.1007/s11999-016-4916-y.
15. Holleyman RJ, Critchley RJ, Mason JM, Jameson SS, Reed MR, Malviya A. Ceramic bearings are associated with a significantly reduced revision rate in primary hip arthroplasty: An analysis from the National Joint Registry for England, Wales, Northern Ireland, and the Isle of Man. *J Arthroplasty*. 2021;36(10):3498-3506. doi:10.1016/j.arth.2021.05.027.