# Recommendations for design and conduct of preclinical *in vivo* studies of orthopedic device-related infection

Moriarty TF[1], Harris LG[2], Mooney RA[3], Wenke JC[4], Riool M[5], Zaat SAJ[5], Moter A[6], TP Schaer[7], N Khanna[8], R Kuehl[8], Alt V[9], Montali A[10], Liu J[11], Zeiter S[1], Busscher HJ[12], Grainger DW[13], Richards RG[1].

[1]AO Research Institute Davos, Clavadelerstrasse 8, 7270 Davos Platz, Switzerland.
[2]Microbiology and Infectious Diseases, Institute of Life Science, Swansea University Medical School, Swansea, UK.
[3]Department of Pathology and Laboratory Medicine, University of Rochester Medical Center, Rochester, NY, US.
[4]Extremity Trauma and Regenerative Medicine Task Area, US Army Institute of Surgical Research, JBSA-Fort Sam Houston, TX, US.
[5]Amsterdam UMC, University of Amsterdam, Dept. of Medical Microbiology, Amsterdam Infection and Immunity Institute, Amsterdam, The Netherlands.
[6]Institute of Microbiology and Infection Immunology, Charité Universitätsmedizin Berlin, Berlin, Germany.
[7]University of Pennsylvania, Department of Clinical Studies New Bolton Center, Kennett Square, PA, USA.
[8]Infection Biology Laboratory, Department of Biomedicine, University and University Hospital of Basel, Switzerland & Division of Infectious Diseases and Hospital Epidemiology, University Hospital of Basel, Switzerland.
[9] Department of Trauma, Hand and Reconstructive Surgery, University Hospital Giessen-Marburg, GmbH, Campus Giessen, Germany
[10] DePuySynthes, Oberdorf, Switzerland.
[11]Tianjin Key Laboratory of Radiation Medicine and Molecular Nuclear Medicine, Institute of Radiation Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College. Tianjin, P.R. China.
[12]University of Groningen and University Medical Center Groningen, Department of Biomedical Engineering, Groningen, The Netherlands.
[13]Department of Bioengineering, and Department of Pharmaceutics and Pharmaceutical Chemistry, University of Utah, Salt Lake City, UT, USA.

*Corresponding author:

T F Moriarty, AO Research Institute Davos, Clavadelerstrasse 8, 7270 Davos Platz, Switzerland. Phone: +41 81 414 2397, E-mail: fintan.moriarty@aofoundation.org

Author contribution statement: All authors contributed their research and/or clinical expertise during the writing and editing of the manuscript, and all authors have approved the submission of the final version.

## Abstract

Orthopedic device-related infection (ODRI), including both fracture-related infection (FRI) and periprosthetic joint infection (PJI), remain amongst the most challenging complications in orthopedic and musculoskeletal trauma surgery. ODRI has been convincingly shown to delay healing, worsen functional outcome and incur significant socio-economic costs. To address this clinical problem, ever more sophisticated technologies targeting the prevention and/or treatment of ODRI are being developed and tested *in vitro* and *in vivo*. Amongst the most commonly described innovations are antimicrobial-coated orthopedic devices, antimicrobial-loaded bone cements and void fillers, and dual osteo-inductive/antimicrobial biomaterials. Unfortunately, translation of these technologies to the clinic has been limited, at least partially due to the challenging and still evolving regulatory environment for antimicrobial drug-device combination products, and a lack of clarity in the burden of proof required in preclinical studies.

Preclinical *in vivo* testing (i.e. animal studies) represents a critical phase of the multidisciplinary effort to design, produce and reliably test both safety and efficacy of any new antimicrobial device. Nonetheless, current *in vivo* testing protocols, procedures, models and assessments are highly disparate, irregularly conducted and reported, and without standardization and validation. The purpose of the present opinion piece is to discuss best practices in preclinical *in vivo* testing of antimicrobial interventions targeting ODRI. By sharing these experience-driven views, we aim to aid others in conducting such studies both for fundamental biomedical research, but also for regulatory and clinical evaluation.

**Keywords:** preclinical study; *in vivo;* biofilm; orthopedic device-related infection; antimicrobial device.

## 1. Introduction and background

The placement of an orthopedic device into patients creates a predisposition and increased susceptibility to fracture-related infection (FRI) and periprosthetic joint infection (PJI), which we collectively term orthopedic device-related infection (ODRI). The two main underlying reasons for this are reported to be: 1) the ability of bacteria to adhere to, and form biofilm on or around the device, and 2) a deficit in host immunological defenses adjacent to the device. Both factors exist for all implanted devices, and so the risk of infection is considered universal for all implanted orthopedic devices. Differences in the device size, anatomical location, incision or wound size, underlying co-morbidities and perhaps predisposing genetic factors largely account for the variability in infection rates for specific device classes.

Bacterial biofilm/microcolony formation is causally central to ODRI (Figure 1) as it provides contaminating bacteria protection from host phagocytes and promotes tolerance to antibiotics. The reduced metabolic activity of bacteria deep in the biofilm, coupled to potential limited penetration of some antibiotics through the biofilm matrix, severely hinders the inhibitory or bactericidal activity of many antibiotics (reviewed in refs. [1; 2]). Additionally, metabolically senescent bacterial populations in biofilms, often called persister cells, are naturally more tolerant to antibiotics [3], and are considered critical in the high recurrence rate of ODRI. The combined presence of a foreign body and bacteria prompts dysregulation of the local immune response [4], over and above the reduced phagocytic activity of host endogenous neutrophils in the presence of foreign materials. These concepts were first characterized by Zimmerli and Gristina in the 1980s [5-7]. Indicative of this phenomenon, tissue beds adjacent to experimentally-implanted devices have been shown to harbor viable microcolonies of pathogenic bacteria for prolonged periods [8], including inside macrophages after phagocytosis [9] More recently, bacteria have been shown to also reside within bone canaliculi, which may be a crucial factor in the failure of treatment without extensive debridement [10].

Considering the risks and mechanisms of infection development around implanted orthopedic devices, coupled to an increasing number of surgically placed orthopedic devices expected globally over the coming decades [11], there is a clear need for improved technologies to prevent, diagnose and treat ODRI [2; 12]. The established pathway to clinical implementation for any anti-infective

technology proceeds from preclinical (*in vitro* and *in vivo*) testing followed by clinical evaluation. As increasing numbers of publications emerge evaluating antimicrobial devices, and as the regulatory requirements for preclinical evaluation of antimicrobial technologies evolve, a notable variety of approaches are taken to demonstrate preclinical efficacy of these technologies[13; 14]. With recent reviews describing best practices in testing antimicrobial technologies focusing upon *in vitro* approaches [15], the need exists for analogous guidance in standardizing preclinical *in vivo* models of ODRI. The recommendations herein are assembled by experienced researchers in the field of *in vivo* evaluation of ODRI-targeting technologies, sharing their expert opinion on best practice in areas such as model selection, study design, data interpretation and targets for efficacy. Where possible, minimum criteria are indicated that we consider mandatory for all testing, in addition to other features that may be desirable or appropriate for only certain subsets of studies.

## 2. Prerequisites to *in vivo* studies

Since inducing an ODRI has potential for imposing a significant burden on the experimental animal, *in vivo* studies should only be performed when certain prerequisites are established, a thorough and comprehensive literature search has been conducted, and supportive *in vitro* studies are either completed or shown to be inadequate to address the research question at hand. A comprehensive review of *in vitro* testing methodologies has recently been published [15], to serve as a useful roadmap for testing of antimicrobial technologies prior to considering an *in vivo* study. A shortlist of *in vitro* testing recommended prior to a preclinical *in vivo* study is shown in Table 1.

A legal requirement prior to commencing preclinical *in vivo* testing is, of course, ethical approval, details of which must be reported upon publication. Any institutional review board (IRB) or institutional animal care and use committee (IACUC) will demand a clearly described hypothesis, with a justification for the study and the chosen animal species/model, accompanied by an appropriate power analysis for sample size. The details for animal husbandry, the surgical procedures, anesthesia, analgesia, perioperative monitoring protocols by qualified personnel and clearly defined humane endpoints are equally important for a robust animal welfare protocol. For many studies in the field, appropriate control groups are mandated, including positive and negative controls or an already commercially available product. Russell and Burch's 3 R's (Reduce, Refine and Replace) remain imperative for improving laboratory animal treatment and utilization

while ensuring scientific quality and rigor in research using animals [16]. A relatively recent initiative, the PREPARE guidelines [17], are a useful resource to ensure adequate practices and planning. Their application to the study design should help gain ethical approval. Importantly, the gain of knowledge should justify the expected burden of the involved animals and consider the 3R's principles [16].

## 3. Rationale supporting animal experimentation

Despite abundant reports of *in vitro* evidence for antimicrobial efficacy, comparatively few *in vitro* results have been shown to translate reliably to *in vivo* antimicrobial efficacy [18]. Numerous examples exist of how *in vitro* data can be misleading when it comes to *in vivo* testing, which have been reviewed elsewhere [14; 18]. Some prominent examples include biofilm formation, whereby *in vitro* grown biofilms in various culture media alter microbial growth kinetics and lack the host derived proteins which may account for much of the biomass of a biofilm *in vivo.* The addition of host proteins to *in vitro* systems addresses this issue to a certain degree, however, the complexity of the *in vivo* situation is not entirely replicated. In fact, biofilm forming ability *in vitro* has never been confirmed to match biofilm formation *in vivo.*

*In vivo* trials have the benefit of allowing the evaluation of novel candidate antimicrobial technologies in controlled *in vivo* conditions that isolate the effect of the investigated technology from any number of potential confounding factors expected to be present in any actual human patient population. Admittedly, not all *in vivo* experimental conditions are comparable, and intrinsic differences between anatomy, physiology and immune systems in humans and laboratory animals are a recognized limitation that should not be underestimated in translation. The natural resistance of laboratory animals to clinically relevant human pathogens remains incompletely understood and represents a weak point in the interpretation of interventional studies targeting ODRI. For example, it can be difficult to establish an infection in rodents since they are far less susceptible to clinically relevant human pathogens like *S. aureus*. However, this is strain-dependent and requires further study to understand if and why this is a true phenomenon.

Preclinical *in vivo* testing is nevertheless demanded by regulatory bodies for safety and efficacy evaluation of antimicrobial devices prior to any possible applications for trial use in humans (e.g.

General Considerations for Animal Studies for Medical Devices, USFDA). ISO guidelines also exist for certain aspects of these studies (ISO 10993-2: Biological evaluation of medical devices Part 2: Animal welfare requirements). The benefits and limitations of preclinical *in vivo* models of infection are summarized in Table 2.

## 4. Key aspects in model selection, animal welfare reporting and study design

A clear vision for the intended clinical application of the new anti-infective technology is critical in determining the appropriate test pathway. One of the most basic design questions, which should be confirmed at the outset, is whether the device is intended to support the prevention of infection (prophylaxis), or alternatively, the treatment of infection (therapy). In many respects, the prevention of infection and the treatment of established infection are significantly different, (Table 3), and this will largely dictate many aspects of the ideal preclinical evaluation and data sought from such studies. Against a background of relatively low infection rates for elective procedures, many interventions target the prevention of infection in all patients to make a compelling business case for development. Establishing prophylactic efficacy, however, in a clinical trial may be a prohibitively expensive proposition for the traditional hardware manufacturers active in the orthopedic space. This is because of the high numbers of enrolled patients required to show a reduction in infection from a low baseline of approximately 1% infection rate. This dilemma remains unresolved at the present time and has contributed to the failure to translate many of the large number of scientific advances regularly seen in the scientific literature[14].

There is a life-long risk for late-developing infections, as may occur via a hematogenous route. To prevent such infections, an anti-infective technology would need to retain activity over the entire life-time of the implant. To our knowledge, no animal ODRI model has recapitulated this scenario, although it would represent a valuable model to determine the role of, for example, antibiotic prophylaxis prior to dental procedures.

In the treatment of an established infection, it is also challenging to identify the appropriate time to allow an infection to develop prior to treatment. The clinical view may be that the ideal experimental infection model should produce clinical symptoms of infection similar to the targeted clinical situation. These symptoms can range from pain, weight loss and radiographic loosening of

the implant, to redness, swelling or drainage of the incisional site. The decision on how long to let an infection develop prior to intervention will likely be a balance between animal welfare and clinical symptoms, and should be judged on a case-by-case basis knowing that the more chronic the infection, the more convincing (but also the more challenging) will be the intervention that successfully resolves that infection. A standard approach towards treatment will in most cases compare current gold standard with and without the novel intervention. In many cases in ODRI, this may involve parenteral antibiotics, unless the intervention explicitly aims to replace parenteral antibiotic therapy. It is important to note that euthanasia of animals still receiving systemic or local antimicrobials should be done with caution as false negatives or temporarily reduced bacterial burden may result due to post-mortem activity of residual antimicrobials (i.e. antibiotic carry-over effects). An antibiotic-free washout period prior to euthanasia will reduce these concerns, although the potential rebound in bacterial growth may also mask a sub-lethal effect of the material. Nevertheless, the risk of false negative results should outweigh the concern for a rebound effect in most cases.

**Model complexity**

Recent reviews describe the number of variables involved in selecting appropriate orthopedic bone repair preclinical models, including species, bone macro- and micro-structure, bone composition and remodeling, and practical cost, husbandry, species and handling issues [19; 20]. Adding infection variables into this experimental preclinical matrix produces substantial complexity. A wide range of *in vivo* ODRI models are also available (reviewed in [21; 22]), ranging from the comparatively simple models amenable to most institutions and facilities [23; 24], to the more complex models requiring more advanced expertise and facilities [25-27] and finally the most complex models where the clinical condition is recapitulated to the closest extent possible [28-30]. Figure 2 outlines the types of animal models available, arranged according to complexity and appropriateness to different stages of product development.

For clinical translation, models that more closely match the specific clinical condition are clearly preferred and increasingly requested by regulatory bodies. However, the need for such complex models is less clear at the proof of concept stage and may be unnecessary for many basic science studies. In all cases, there remains a necessity to properly justify an *in vivo* experiment, and the

chosen model for its relevance in addressing specific translational questions: animal models should never be used when they cannot effectively address a significant hypothesis. Certain situations, however, may demand a complex model to fully assess the device even in a basic science study, particularly when the function of that device can influence the biology of the surrounding tissues. For example, a technology to prevent infection in patients receiving a fracture fixation device may need to be tested in a model that includes a fracture or soft tissue damage, since biomechanical forces and soft tissue damage influence the healing response of bone, which can also influence host antibacterial defenses [31; 32]. Similarly, fracture models will also allow estimation of the impact of the technology on fracture healing in addition to the antibacterial efficacy, which is also a key outcome for both basic science and regulatory approval.

In certain cases, the clinical situation may not be reliably or easily replicated in any existing model due to practical or animal welfare reasons such as duplicating certain co-morbidities specific to humans, osteoporotic conditions, chronic infected burn wounds or diabetic foot osteomyelitis. Of note, certain clinical conditions may also occur naturally in veterinary clinical practice (e.g. PJI, or infected non-union/pin tract infection in pets) and could offer an avenue for efficacy assessment in future [33; 34].

In general, animal model selection during product development is driven by the overarching objective in the preclinical space of de-risking emerging technologies. During the discovery phase, animal models often help explore and answer hypothesis-driven aims addressing mechanistic concepts. When an early stage technology aligns with an unmet clinical need in the target patient, de-risking becomes a priority. De-risking is driven by a plethora of factors and spans the gamut from economic viability, intellectual property, manufacturing and preclinical safety and efficacy. The stage of product development usually drives the choice of preclinical model. For example, generic safety profiles of candidate test articles can be established cost-effectively in standard murine models. As the path of preclinical development advances, the value proposition of the technology increases, and the clinical relevance of the animal model is of increasing importance. For regulatory purposes in the validation of any new technology, the chosen model should ideally consider the targeted human patient population and reflect the value-proposition for target patients, although this can of course be difficult to achieve or even measure. For orthopedic devices

in general (not specifically anti-infective devices), the selection of a model may also be influenced by key safety and performance aspects such as implant failure modes, duration of the evaluation period, device size, the intended use, and the intended regulatory strategy[35]. This can include consideration of implant dimensions and the surgical approach in the animal compared to standard human use. For regulatory submissions, the implantation of the actual device as intended for human clinical use may in fact be required. Sheep or swine models are therefore appropriate orthopedic animal model choices in this regard, based on implant size alone. Most safety testing for mechanical devices (e.g. joint and valve replacements) required for FDA IDE/PMA regulatory purposes is therefore conducted in relatively large animals, such as rabbits, dogs, pigs, goats, sheep, cattle and large non-human primates [20; 36].

Some human devices are simply too large for application in certain species, and so the ability to produce and use a smaller, analogous "copy" at reasonable cost, and still recapitulate the human situation is important. With functional fracture fixation implant systems available for mid-size animals such as rabbits[37] and even smaller animals such as rats[38] and mice[31; 39], definitive studies may focus on the model with most clinical relevance, rather than on size alone. It should be noted, however, that many types of medical devices, including total joint prostheses, are extremely difficult or impractical to miniaturize or apply in animals within reasonable burden limitations. In addition, the miniaturized device may exhibit significantly altered behavior (e.g. tissue physiology, failure modes) than the human-sized device. Testing of the functional implant may therefore not be suited to small animals for many devices.

Recently, the scientific community and select industry ventures started to embrace naturally occurring disease models in both dogs and horses with the rationale that their use demonstrates robust superiority over experimentally induced models [40]. The latest efforts of refinement and optimization of large animal models including genetically modified models is a positive development towards improving reproducibility and scientific rigor. While there is still paucity of available model data pertaining specifically to ODRI, the authors encourage investigators to stay current in the field as more refined models are being developed and validated [41].

**Reporting animal care and use**

*In vivo* ODRI studies exert increased burden upon test animals over and above non-inoculated equivalents. Animal welfare practices are, therefore, important to fully address and disclose when publishing. A relatively recent initiative, the ARRIVE guidelines [42], provides a useful resource to ensure adequate reporting. Minimum requirements: all baseline information such as animal species/strain, gender, age, weight, immune status, microbiological status (e.g. specific pathogen-free (SPF), or specific and opportunistic pathogen-free (SOPF)), overall husbandry, detailed description of the surgical intervention and veterinary care, including full anesthesia and analgesia protocols often known to affect animal health status and recovery should be reported, or made available through the primary description of the chosen model. The observation parameters used, including contingency plans for addressing possible complications, and the results thereof, should also be reported.

A score sheet is a useful means to objectively report the observation of animals based on clinical parameters, but also to define when an animal should be excluded from the study (i.e. humane endpoints) and may be adapted to each specific study [43] (Table 4). Frequency of scoring should be defined according to the study phase and must increase if the clinical condition warrants it. The full and accurate reporting of animal welfare issues is useful with increasing welfare demands on the animals for a given model, but also helps predict or anticipate losses for others utilizing the model. The clinical parameters and observations that may be considered as minimum requirements include: vitals, weight loss, wound healing issues, behavioral changes, and any animals excluded from the study, including the reason for exclusion (e.g. severe symptoms, found dead, sepsis). Adequate pain management is also critical to disclose, including any pain medication administered in addition to pre-planned pain control.

**Pilot studies, validation of ODRI models and historical controls**

The reliability and scientific rigor of the experimental outcomes in preclinical *in vivo* studies will be improved by using established animal models with known responses or behaviors under standardized and controlled conditions. Whenever using an animal model for the first time, developing a new model, or adapting an existing model (e.g. applying a different pathogen, inoculation dose, time to treatment, or use of a systemic antibiotic), pilot studies are mandated to establish baseline infection rates as a *minimum* requirement. Pilot studies allow the investigator

to refine some of the procedures within the study and approximate the effect size to determine the sample size for a pivotal study. This can result in less discomfort for the animals and may aid in the reduction of the overall number of animals required.

Certain additional prior information on the behavior or performance of a model would be desirable, though perhaps less critical to provide a complete understanding of the model. For example, it would be valuable to know whether the original infection persists in the model of choice or whether it is self-resolving. This is important because a number of infection models have been described to be self-resolving, mainly for infections with low virulence organisms or with high tolerance (rodent) hosts [44] and given a certain duration. In such cases, the time point chosen as an endpoint becomes crucial to properly evaluate results. Similarly, it is valuable to know whether an infection develops or persists under conditions of conventional clinical prophylaxis or treatment. If conventional prophylaxis is adequate in a certain model, the novel technology may require an additional proven benefit, such as reduced systemic toxicity or ease of application, to produce a clear benefit to the patient.

**Disease state and comorbidities**

Clinically, the risk of developing ODRI and the severity of such an infection are impacted by the patient's co-morbidities and it may be a rational strategy to attempt to include these co-morbidities in an animal model for certain studies. Among the most recognized of these co-morbidities are obesity, diabetes, immunosuppression, and coagulopathy [45]. Considering the frequency of these co-morbidities in orthopedic surgeries, understanding the contribution of these factors to infection initiation and progression is critical to ameliorating ORDI risk during and after surgery.

Rodent models have been most frequently employed to study the impact of obesity and diabetes on infection. High fat diets (HFD) containing 40% or greater fat content are fed to rodents for 2-4 months to promote increases in adiposity and glucose intolerance. Male C57BL/6 mice have traditionally been used since they consistently become obese (i.e. greater than 30% increase in body mass compared to low fat diet controls) and develop metabolic syndrome (i.e. insulin resistance and glucose intolerance) on these diets. Female C57BL/6 mice have less consistent response to the HFD and are used less often. However, most but not all commonly available strains

also show weight gain and metabolic syndrome on a HFD [46]. Choice of mouse strain to employ may be dictated by consistency with prior investigations, in which case C57BL/6 is the likely choice, or by consideration of other genetic characteristics of the strain (see below). Morbid obesity is a dominant co-morbidity. Leptin deficient (*ob/ob*) and leptin receptor deficient (*db/db*) mice are classic models for this clinical state. Unfortunately, the disruption of leptin function has clearly documented effects on immune function, including risk of infection [47; 48]. Thus, these genetic models are inappropriate for studying the mechanisms of obesity-mediated ODRI. Other studies have used obese Zucker rats in bone healing and orthopedic trauma models [49; 50]. The obese Zucker rat, homozygous for the *fa* allele, spontaneously becomes obese and is an accepted model of metabolic syndrome, sharing many similarities with humans who have this condition, including obesity, dyslipidemia, some insulin resistance, and hypertriglyceridemia. No implant infection studies are yet reported with this model.

While a suppressed immune system is an obvious co-morbidity leading to increased risk for ODRI, more subtle differences in the immune state are an important consideration when selecting a rodent model to study infection. Recent evidence suggests that the pro-inflammatory Th1/Th17 response to infection may be more effective in the early planktonic growth phase of an infection and comparatively less so in the chronic biofilm stage [51]. In the chronic biofilm stage, the anti-inflammatory Th2/Treg immune response seem to be more effective. For example, C57BL/6 mice show a Th1/Th17-biased response to infection while BALB/c mice show a Th2 and Treg bias. Consistent with this logic, BALB/c mice have been shown to clear ODRI biofilm infection caused by *Staphylococcus aureus* more effectively than C57BL/6 mice [51]. While these assertions must be confirmed in future studies, the Th1/Th2 balance clearly has relevance in selecting certain animal species and strains for ODRI.

Infection and its associated inflammation lead to activation of the coagulation pathways including down-regulation of anticoagulation mechanisms and inhibition of fibrinolysis [52]. Many aspects of the coagulation pathways can be interrogated for their contribution to risk or progression of infection through genetic mouse models. For example, in a mouse model of *S. aureus* ODRI, transgenic overexpression of plasminogen activator inhibitor-1 (PAI-1) led to markedly more abscess communities (SACs) and more intense fibrin encapsulation of these abscesses [53]. Co-

morbidity studies emulating this example provide evidence that the coagulopathy of infection can play a direct role in the infection process. Additional infection studies using genetic disruption of the coagulation cascades with implant infections are warranted.

**Outcome measures**

In most studies, the primary outcome is the reduction of infection rate and/or bacteria, but may also be changes in immune responses, or parameters addressing functional restoration i.e. improvements in fracture healing, bone regeneration or osseointegration. Clinical signs of infection such as white cell count, C-reactive protein, swelling, weight loss or lameness are other measures useful to evaluate outcome. However, since the intervention may reduce these symptoms in culture-positive animals, clinical signs should only be used as a secondary outcome measure supporting the more definitive outcome measure provided by quantitative bacterial culture.

From a translational standpoint and a value-based health care system, it is important to understand how preclinical results de-risk and support the value-proposition for the target patient. From a clinical perspective, complete eradication of infection from the implant and surrounding tissues and restoration of function is the ideal target for most anti-infective technologies, the easiest result to interpret, and likely greatest impact in terms of asserting efficacy. In contrast, the clinical significance of, for example, a 90% reduction in bacteria is a challenge to interpret, regardless of any statistical differences proven. *In vitro* studies of antibacterial activity often target a 1'000-fold reduction in bacterial as a significant, yet arbitrary threshold. Also, a 1'000-fold reduction from a reference point of $10^7$ CFU may have different significance in terms of outcome compared with a 1'00-fold reduction from $10^4$ CFU. At the present time, no generally accepted target has been established for a CFU reduction in *in vivo* studies. Clearly, complete eradication of infection in all test-treated animals represents an ideal target for many antimicrobial-containing devices, and so should be aimed for from the outset of any study. Any remnant bacteria in a technology "merely" reducing the bacterial load may have an impact, but it remains likely that remaining bacteria may proceed to re-establish a biofilm-related infection and so is not considered a complete cure.

Histological evaluation of outcome is an important assessment, and when performed thoroughly serves as a useful adjunct to culture results, because it may pick up inflammation indicative of infection when culture results remain negative. Identifying an animal as infected or not based solely on histological sections is to be undertaken with caution due to the limited sample areas in traditional histology. A technique combining histology with molecular detection of pathogens is fluorescence *in situ* hybridization (FISH), which is also a marker for bacterial activity [54] (Figure 3). Conventional histopathology also has a valuable role in monitoring ODRI in *in vivo* studies. Some features of *S. aureus* infection, such as SACs, as well as peri-implant osteolysis and inflammatory cell infiltration are best revealed with conventional histopathology[55-57].

**Study design and outcome assessment**

In performing an *in vivo* study, inherent bias, either intentional or unintentional, should be avoided [58]. *As* a minimum, randomization should always be performed, as well as blinding or masking the study personnel to groups/cohorts and treatments when possible. Surgical procedures and certain therapies (e.g. coated implants) are difficult to blind compared to drug therapy, for which a placebo is also provided as control. In such circumstances, an empty coating or delivery vehicle may serve as a suitable control. Under circumstances where blinding of the operator (e.g. surgeon) is impossible, de-blinding should be done at the latest timepoint possible (e.g. just before treatment application) and study design should be masked to the person performing outcome measurements (e.g. biopsy culturing, histology).

Efficacy must also be compared to appropriate positive and negative controls. In many cases, there may not be an apparent positive clinical control or defined standard of care to compare the experimental therapy. This is especially true for more innovative therapeutic strategies. In such cases, conventional interventions such as debridement and systemic antibiotic may be the best option for a positive control. A negative control for ODRI may involve debridement alone or implant removal. In an established animal model of infection, historical controls may have some value, however, it is highly recommended to include a limited number of additional control animals in each experiment to ensure reproducibility of the model.

With regards to statistical evaluation, categorical data (e.g. infected vs. not infected) assessed using, for example, Fisher's Exact Test or Chi Square need robust differences between groups or

large sample sizes. Additionally, more subtle differences may exist in the levels of infection. Because of high biological variability in CFU counts, these values often are not normally distributed, and non-parametrical statistics are indicated. The need for such clear differences between infected and non-infected animals, is, in part, why many investigators compare mean/median viable counts to determine differences between groups. For a general review on data transformation and statistical comparison in *in vivo* studies, please see the following reviews [59; 60].

## 5. Bacteriological aspects of ODRI studies

### Pathogen selection

The clinical manifestation of ODRI may vary depending upon the causative microorganism itself and not just the host immune response or device type in question. The infecting pathogen, its antibiotic susceptibility and resistance profiles, ability to form biofilm and repertoire of virulence factors possessed by that species/strain, are therefore a critical feature of an *in vivo* study. The most commonly reported pathogens reported in clinical studies of ODRI include *S. aureus* (MSSA and MRSA), Coagulase negative Staphylococci (including *S. epidermidis*, Enterococci, *Pseudomonas aeruginosa,* Enterobacteriaceae, anaerobes and aso include a significant portion of polymicrobial infections. It is important to note that more virulent species such as *S. aureus*, and aerobic Gram-negative bacilli are the major causes of early onset infections, whilst delayed or late infections are commonly caused by less virulent species such as coagulase-negative staphylococci (CoNS, e.g. *Staphylococcus epidermidis*) [61; 62].

With regards to specific strain selection, molecular epidemiological studies have begun to reveal the full extent of genetic diversity within *S. aureus* and *S. epidermidis* populations [63]. In the past, *in vitro* and *in vivo* studies have used American Type Culture Collection (ATCC) or other well-characterized lab strains such as *S. aureus* ATCC 33591, *S. aureus* ATCC 49230 (isolated from a patient with chronic osteomyelitis), *S. aureus* 8325-4, *S. aureus ATCC* 43300*, S. aureus* Newman, *S. epidermidis* ATCC 35984 (RP62A), and *S. epidermidis* ATCC 35983 (RP12); or unique or random clinical isolates such as *S. aureus* MN8 [64; 65] and *S. epidermidis* 1457 [66; 67]. At the present time, there appears to be no consensus as to the best strain or rationale for strain selection. Culture collection strains have value in international access and traceability, and quite often a significant amount of

characterization including sequencing and performance in *in vivo* models. However, there is concern that even clinical isolates deposited in a culture collection may have lost certain virulence capabilities, or at least this is a challenge to quantify or certify. Freshly isolated clinical isolates therefore have an advantage of being closest to the active pathogen when in humans. However, these are lacking in accessibility and comparability to other models and may also lose virulence once passaged and distributed internationally. In the absence of clear guidance, it appears that the ideal situation would be to have a basic understanding of the clinical background to the strain, match it to the clinical question being modelled, and to know key genetically encoded virulence factors retained by the bacterium. In case more than one strain is analyzed, considering the use of particular "benchmark" strains (still to be selected) might provide better inter-study comparisons. A key virulence factor of *S. aureus* and in particular *S. epidermidis* is the ability to adhere and form biofilms directly on implant surfaces, which then protects them from antibiotics and the host's immune system [68; 69]. Both *S. aureus* and *S. epidermidis* are known to use cell-wall-anchored (CWA) proteins/adhesins such as fibrinogen (ClfA/Fbe/SdrG), fibronectin (FnBP/Ebh/Embp), collagen (Cna/SdrF), vitronectin (VnBP/AtlE,/Aae), bone sialoprotein (SdrE/Bhp) and elastin (FnbpA/EbpS) to initially adhere to the implant surface and to host extracellular matrix that covers the implant [70-74]. They then accumulate to form a biofilm using various independent mechanisms, namely the polysaccharide intercellular adhesin (PNAG/PIA), synthesized by *icaADBC*-encoded proteins [66; 75; 76] or by proteinaceous factors independent of the *icaADBC* locus, which include *S. aureus* cell wall protein G (SasG) and *S. epidermidis* accumulation associated protein (Aap) [77; 78]; extracellular matrix-binding protein (Ebh/Embp) [79-81]; biofilm-associated protein (Bap) [82]; and extracellular DNA (eDNA) [83; 84]. Thus, the fact that closely related isolates can display different phenotypes and produce the above proteins differently and influence the ability of the *S. aureus*/*S. epidermidis* to adhere and form biofilms is an important factor to consider when choosing an isolate for certain *in vivo* studies. An *in vivo* ODRI study should therefore consider including a characterized benchmark strain of *S. aureus* and/or *S. epidermidis* known to form a biofilm by PNAG/PIA and SasG/Aap and Aap/SasG production, such as *S. aureus* MN8, *S. aureus* BAA-1707 (MW2), *S. epidermidis* ATCC 35984 (RP62A), *S. epidermidis* 1457 and *S. epidermidis* 5179R1. A good summary of the characteristics of different types of *S. aureus* isolates, specifically methicillin-resistant *S. aureus* (MRSA) strains, can be found in the supplementary data of Moneke *et. al*. [85].

Antibiotic resistance is another factor in identifying an appropriate pathogen for an *in vivo* ODRI study. Profiling the susceptibility of the selected pathogen to any antibiotic used in the study is essential. Resistance of biofilms formed by the isolate to the same antibiotic may also be beneficial for many studies, with care taken to report the protocols and methods used completely [86; 87].

The last consideration for the strain/pathogen selection is whether to use fluorescently labelled isolates or bioluminescent isolates that with the proper optical imaging instrumentation can be visualized *in situ* [88; 89]. Stable fluorescent-labelling requires genetic manipulation of the bacteria which is not always readily achievable and is particularly problematic when clinical isolates are being investigated. In any case, the resulting virulence of the modified pathogen must be compared to the unmodified wild-type isolate. Some research groups have now developed a variety of fluorescent reporter plasmids for labelling *S. aureus* by utilizing plasmids encoding either green fluorescent protein (GFP) or higher wavelength reporter variants for yellow (YFP) and red (mCherry) labelling [24; 88; 90-92]. These reporters are placed under control of characterized promoters to enable constitutive or inducible expression using antibiotics such as chloramphenicol, erythromycin or tetracycline [90; 93]. While fluorescent strains are very useful for *in vitro* studies, they have very limited value for *in vivo* studies because of the low intensity of the signal and quenching by the host tissue. In contrast, use of bioluminescent strains in animal models of ODRI can provide a longitudinal outcome measure of *in vivo* bacterial growth[94]. For therapy studies, bioluminescence can also be used to confirm the establishment of infection before randomizing the animals into treatment groups [56] However, *in vitro*, it has been shown that although bioluminescent signals correlate with planktonic growth, this correlation is lost during the transition from planktonic to the biofilm mode of growth [95]. In line with this, any lack of luminescent signal *in vivo*, or reduction of signal over time, does not necessarily indicate reduction of viable numbers of bacteria [96]. Apparently, viable bacteria with low metabolic state may persist and go unnoticed because of reduced luminescence. Thus, *in vivo* investigations of antimicrobial strategies results with bioluminescent strains should be interpreted with caution. Another point of attention is the localization of the infection; deep infections are less likely to be reliably monitored than superficial infections due to the different path length for the emitted light from the bacteria, and (non-physiologically) high numbers of bacteria may be needed to detect a signal.

## Inoculum preparation

While often overlooked or sparsely described, inoculum preparation protocols must be clearly defined, and should be fully disclosed in reporting or publishing animal studies of ODRI. This is important as the infection rate in any model will be related to the inoculum, and the ability to replicate any model across different laboratories must rely on this information. The essential minimum information to share includes the total number of colony forming units (CFU) administered to each animal, but also the exact conditions used for preparation of the inoculum e.g. using freshly prepared bacterial inocula or frozen aliquots, fresh bacteria suspended from colonies on agar plates of harvested from liquid culture, type of culture media, log phase or stationary phase bacteria, growth media, application of inocula in saline solution, pre-colonized on implants, or deposited with a collagen sponge, fibrin gel or other carrier material. Similarly, it is important to describe the storage conditions, including the time between preparation and administration of the bacterial inoculum, in addition to the means used to apply/deliver the bacteria into the target tissues, the volume and any method to limit the spread of bacteria in the implant or tissue site.

## Diagnosis/definition of implant related infection

Preclinical *in vivo* studies by their nature, offer the possibility for thorough bacteriological and histological evaluations of the experimental animal. Basic knowledge of microbiological techniques regarding species identification and quantitative culturing methods is a prerequisite for reliable results. Since normal biological variation in animal experiments is a known, intrinsic challenge, attention should be directed to standardized sampling and homogenization or sonication of all samples to obtain reliable CFU counts. Similarly, all sampling locations should be accurately described, recognizing that small biopsies are at risk for missing bacteria; ideally each animal may be sampled in its entirety and include both bone and soft tissue as well as the implant (after sonication). For most models, sampling is only possible at a revision surgery or post mortem. However, the tissue cage model [23], does allow repeat sampling of the tissue cage fluid, which has significant advantages in terms of continuous monitoring of bacterial burden, immune cells or extracellular immune mediators.

Detection of infecting pathogens on the implant or in the tissues by conventional culture will usually have high sensitivity and specificity. However, clinical literature documents culture-negative infections. This occurs when biopsies are culture negative, either due to sample error (not necessarily an issue in animal studies) or due to non-cultivable bacteria in some situations[97; 98]. Small colony variants (SCVs), subpopulations with a slow growth rate and atypical colony morphology, may develop and can only be detected in prolonged cultures [99]. According to the current hypothesis, bacteria may also enter a viable but non-culturable (VBNC) state in particular within biofilms, but also as planktonic bacteria and under antibiotic treatment. This VNBC state may render bacteria unable to regrow when deposited on conventional microbiological media [98]. This issue may be more prevalent for certain bacterial species, or for patients/animals undergoing antibiotic therapy. However, the exact conditions and mechanisms underlying the VNBC concept remain to be fully understood, but do represent a possible risk in ODRI studies. To maximize chances of accurate culture results, adopting clinical methods towards improved isolation and culture of biopsies should also be considered, i.e. sonication of biopsies and associated hardware components [100; 101] or the use of known biofilm-dispersing agents during sample processing[102].

Molecular technologies (e.g. PCR) that amplify bacterial DNA in samples have improved the culture- independent detection and identification of microorganisms in the past years [103]. These techniques can be very sensitive and fast but must be thoroughly evaluated for the respective experimental setting. An efficient DNA-extraction protocol is crucial, and results are often only semi-quantitative, because the background of eukaryotic DNA in samples may hamper detection of the relatively minute amount bacterial DNA. In addition, these techniques detect bacterial DNA, but cannot indicate if the bacteria were living or dead at the time of sampling [104].

Since culture as well as amplification-based methods require disintegration of the sample (and therefore disruption of biofilms), microscopic techniques are the only methods to date that can differentiate between the presence of single cells, microcolonies, and biofilms in tissues [105]. In addition to Gram staining or immunohistochemical methods, FISH combines molecular detection of microorganisms with fluorescence microscopy and has been increasingly used for analysis of biofilm-associated infections [103; 106; 107]. FISH can be applied to *in vitro* samples as well as to *ex vivo* samples from animal models or patients [108-111]. Since FISH-probes hybridize ribosomal RNA, the

signal intensity correlates to the ribosome content and consequently activity of the cells. This allows both visualization and identification of the microorganisms, and also provides information about their spatial distribution and activity *in situ* [112] (Figure 4). FISH can localize and quantify planktonic cells, biofilms, or intracellular bacteria and can clarify potential problems such as contamination and mixed infections, or inoculation in an incorrect compartment (Figure 4). Consequently, FISH is increasingly used in infection models [109; 111; 113], although it is labor intensive and rather restricted to specialized laboratories.

Finally, all bacteriological evaluations should involve a method to confirm that the infection observed *in vivo* is caused specifically by the inoculated strain and not another opportunistic host pathogen. In some cases, this may be performed on simple selective microbiological agars and subsequent phenotypical and/or biochemical characterization. More advanced PCR-based or sequencing techniques (e.g. Random Amplified Polymorphic DNA (RAPD)) will not only enable confirmation of the species identity of the infecting isolate, but also the strain and whether it matches the inoculated strain. Using fluorescently labeled bacterial strains might be a simpler way to confirm the identity of the inoculated strain, although co-infecting pathogens should also be ruled out (e.g. [24]).

## 6. Antibiotic/Antimicrobial administration

In most preclinical *in vivo* studies, antimicrobial agents are often administered, either as part of the technology being tested, or as conventional administration as an adjunctive or comparison therapy. In general, minimal evaluation of the antimicrobial in use may require estimating the loading of the drug in the carrier, its release from the carrier or implant and its quantification in plasma or local tissues. Accurate, reliable estimates of these values require validated methods, with known limits of detection, limit of quantification, specificity and accuracy in physiological media. Antibiotic concentrations measured in tissue must be carefully considered, with awareness of the distinctive partitioning of different antibiotics into different tissue niches and protein binding [114].

**Systemically applied**

Certain studies may require systemic antibiotic coverage either prophylactically or therapeutically. The goal of systemic antibiotic therapy may be to mimic human standard of care, or to prevent systemic infection. Systemic antibiotic therapy may be particularly appropriate in testing technologies that provide protection of the device, but do not release antimicrobials locally into the surrounding tissues (e.g. contact-killing or anti-adhesive coatings, bound agents). Administration of systemic antibiotics should be supported by a clear, clinically consistent rationale for antibiotic selection, timing, dosage, route and frequency. It is desirable to match the human clinical condition as closely as possible to ensure adequate regimens are used. Of course, safety and pharmacokinetics may need to be established in advance. Clearance of antibiotic from systemic circulation may be different in animals than in humans, and any study with a significant systemic antibiotic component, or the use of antibiotics not approved for animal use should consider a preliminary pharmacokinetic profile analysis. Simultaneously, potential (unreported) side effects can be investigated. Persistence of reliable ORDI in the model using systemic antibiotic regimens alone must first be asserted before efficacy of any other additional implant-focused antimicrobial strategies can be distinguished.

While clinically relevant, no clear recommendations exist for testing an antimicrobial device in the absence versus presence of systemic antibiotic therapy. Decisions may be best dictated by the clinical problem in question, the experience with the animal model, and the expected protective or therapeutic bioactivity of the implant technology.

**Locally applied**

The majority of new antimicrobial technologies applied to medical devices involve antimicrobial agents released from the implant or carrier. As mentioned above, the total amount of antimicrobial dosed to each animal should be clearly described as a minimum, and an estimation of antimicrobial release (or retention) provided to assess the performance of the material/agent *in vivo*. It is also desirable to know the tissue/serum concentration of the released antimicrobial over time as well as a measurement of observed toxicities (both local and systemic) associated with local application of the material or antimicrobial. Local concentrations of antimicrobials in the surrounding tissue may be measured using ultrafiltration[115] or micro-dialysis techniques[116], *post mortem* or in biopsies using tissue extraction and known drug analysis methods[116].

## 7. Common errors in *in vivo* ODRI studies

The earliest reports of using animals to study ODRI date back many decades. In these earliest studies, certain protocols were used that are no longer considered best practice. There is a risk that these older practices persist when these older studies are used as a reference for new studies. Some of the more common errors in these studies are shown in Table 5.

## 8. Outlook for Best Practices

The goal of this opinion piece is to present a minimum set of requirements to be considered when planning an ODRI animal study. Broad application of these experimental design principles would aid the proper execution of preclinical animal ODRI studies so that improved results of these studies carry maximum weight and reliability, can be more easily transferred between laboratories, and can better support translation of these technologies into clinical practice. A list summarizing the key points recommended to be reported when publishing such studies is included as reference for the reader (Table 6). Although many outstanding questions for the field remain, guidance recommendations for best practices here should prompt further discussion in the ORDI research forum hopefully with the intent to standardize and validate these approaches. Importantly, many practices recommended here should also be listed in publications in the field to reinforce their value. Further methodological refinements or more absolute policy statements may be achieved via consensus between scientific and regulatory agencies and with careful reporting of future studies with provisions and details described here as a starting position. Such collaboration between researchers, regulatory frameworks, and medical device industry could also further clarify and maximize what can be reliably and accurately shown in animal studies and hopefully also minimize what must be shown in a clinical setting [13; 14].

Given increasing reports of animal studies investigating ODRI from many geographic regions, steps towards global harmonization of methods, analytical and microbiological approaches, reporting requirements, reliability and reproducibility, validation and translational value must be considered seriously [2]. Experimental facilities capable of performing *in vivo* preclinical studies are available across the globe. Hence, the conditions and resources of a research facility are not often a major barrier to conducting properly designed ODRI preclinical studies to safe, acceptable and

appropriate standards. Establishing minimum standards in designing and reporting these ODRI preclinical studies might stimulate wide-spread adoption and further, promote needed global harmonization and understanding of best practices. International guidelines are already available for governing animal welfare in medical research, and the antimicrobial implant study related parameters suggested in this document now provide a solid additional basis conducting antimicrobial preclinical studies in ODRI.

## 9. Acknowledgements

TFM, MR, NK, RK, AM, LGH and SAJZ would like to acknowledge networking support by the COST Action iPROMEDAI (Project No. TD1305), supported by COST (European Cooperation in Science and Technology). TFM, RAM and RGR were supported by the AOTrauma CPP on Bone infection. Henk J. Busscher is also director of a consulting company, SASA BV (GN Schutterlaan 4, 9797 PC, Thesinge, The Netherlands). The authors declare no potential conflicts of interest with respect to authorship and/or publication of this article. Opinions and assertions contained herein are those of the authors and not construed as necessarily representing views of the funding organization or their respective employees.

## 10. References


1. Lebeaux D, Ghigo JM, Beloin C. 2014. Biofilm-related infections: bridging the gap between clinical management and fundamental aspects of recalcitrance toward antibiotics. Microbiol Mol Biol Rev 78:510-543.
2. Hall-Stoodley L, Stoodley P, Kathju S, et al. 2012. Towards diagnostic guidelines for biofilm-associated infections. FEMS Immunol Med Microbiol 65:127-145.
3. Harms A, Maisonneuve E, Gerdes K. 2016. Mechanisms of bacterial persistence during stress and antibiotic exposure. Science 354.
4. Zaat S, Broekhuizen C, Riool M. 2010. Host tissue as a niche for biomaterial-associated infection. Future Microbiol 5:1149-1151.
5. Zimmerli W, Lew PD, Waldvogel FA. 1984. Pathogenesis of foreign body infection. Evidence for a local granulocyte defect. J Clin Invest 73:1191-1200.
6. Zimmerli W, Waldvogel FA, Vaudaux P, et al. 1982. Pathogenesis of foreign body infection: description and characteristics of an animal model. J Infect Dis 146:487-497.
7. Giridhar G, Gristina AG, Myrvik QN. 1993. Altered oxidative responses and antibacterial activity of adult rabbit alveolar macrophages exposed to poly(methyl methacrylate). Biomaterials 14:609-614.
8. Broekhuizen CA, de Boer L, Schipper K, et al. 2007. Peri-implant tissue is an important niche for Staphylococcus epidermidis in experimental biomaterial-associated infection in mice. Infect Immun 75:1129-1136.
9. Boelens JJ, Dankert J, Murk JL, et al. 2000. Biomaterial-associated persistence of Staphylococcus epidermidis in pericatheter macrophages. J Infect Dis 181:1337-1349.
10. de Mesy Bentley KL, Trombetta R, Nishitani K, et al. 2017. Evidence of Staphylococcus Aureus Deformation, Proliferation, and Migration in Canaliculi of Live Cortical Bone in Murine Models of Osteomyelitis. J Bone Miner Res 32:985-990.
11. Kurtz S, Ong K, Lau E, et al. 2007. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am 89:780-785.
12. Moriarty TF, Kuehl R, Coenye T, et al. 2016. Orthopaedic device-related infection: current and future interventions for improved prevention and treatment. EFORT Open Rev 1:89-99.
13. Busscher HJ, van der Mei HC, Subbiahdoss G, et al. 2012. Biomaterial-associated infection: locating the finish line in the race for the surface. Sci Transl Med 4:153rv110.
14. Grainger DW, van der Mei HC, Jutte PC, et al. 2013. Critical factors in the translation of improved antimicrobial strategies for medical implants and devices. Biomaterials 34:9237-9243.
15. Sjollema J, Zaat SAJ, Fontaine V, et al. 2018. In vitro methods for the evaluation of antimicrobial surface designs. Acta Biomater 70:12-24.
16. Tannenbaum J, Bennett BT. 2015. Russell and Burch's 3Rs then and now: the need for clarity in definition and purpose. J Am Assoc Lab Anim Sci 54:120-132.
17. Smith AJ, Clutton RE, Lilley E, et al. 2018. PREPARE: guidelines for planning animal research and testing. Lab Anim 52:135-141.
18. Moriarty TF, Grainger DW, Richards RG. 2014. Challenges in linking preclinical anti-microbial research strategies with clinical outcomes for device-associated infections. Eur Cell Mater 28:112-128; discussion 128.
19. Auer JA, Goodship A, Arnoczky S, et al. 2007. Refining animal models in fracture research: seeking consensus in optimising both animal welfare and scientific validity for appropriate biomedical use. BMC Musculoskelet Disord 8:72.
20. Pearce AI, Richards RG, Milz S, et al. 2007. Animal models for implant biomaterial research in bone: a review. Eur Cell Mater 13:1-10.
21. Reizner W, Hunter JG, O'Malley NT, et al. 2014. A systematic review of animal models for Staphylococcus aureus osteomyelitis. Eur Cell Mater 27:196-212.
22. Calabro L, Lutton C, Din AFSE, et al. 2013. Animal Models of Orthopedic Implant-Related Infection. In: Moriarty TF, Zaat SAJ, Busscher HJ editors. Biomaterials Associated Infection: Immunological Aspects and Antimicrobial Strategies. New York, NY: Springer New York; pp. 273-304.
23. Nowakowska J, Landmann R, Khanna N. 2014. Foreign Body Infection Models to Study Host-Pathogen Response and Antimicrobial Tolerance of Bacterial Biofilm. Antibiotics (Basel) 3:378-397.

24. Riool M, de Boer L, Jaspers V, et al. 2014. Staphylococcus epidermidis originating from titanium implants infects surrounding tissue and immune cells. Acta Biomater 10:5202-5212.
25. Lucke M, Schmidmaier G, Sadoni S, et al. 2003. Gentamicin coating of metallic implants reduces implant-related osteomyelitis in rats. Bone 32:521-531.
26. Moriarty TF, Debefve L, Boure L, et al. 2009. Influence of material and microtopography on the development of local infection in vivo: experimental investigation in rabbits. Int J Artif Organs 32:663-670.
27. Odekerken JC, Walenkamp GH, Brans BT, et al. 2014. The longitudinal assessment of osteomyelitis development by molecular imaging in a rabbit model. Biomed Res Int 2014:424652.
28. Moriarty TF, Schmid T, Post V, et al. 2017. A large animal model for a failed two-stage revision of intramedullary nail-related infection by methicillin-resistant Staphylococcus aureus. Eur Cell Mater 34:83-98.
29. Stewart S, Barr S, Engiles J, et al. 2012. Vancomycin-modified implant surface inhibits biofilm formation and supports bone-healing in an infected osteotomy model in sheep: a proof-of-concept study. J Bone Joint Surg Am 94:1406-1415.
30. Metsemakers WJ, Schmid T, Zeiter S, et al. 2016. Titanium and steel fracture fixation plates with different surface topographies: Influence on infection rate in a rabbit fracture model. Injury 47:633-639.
31. Sabate Bresco M, O'Mahony L, Zeiter S, et al. 2017. Influence of fracture stability on Staphylococcus epidermidis and Staphylococcus aureus infection in a murine femoral fracture model. Eur Cell Mater 34:321-340.
32. Kalicke T, Schlegel U, Printzen G, et al. 2003. Influence of a standardized closed soft tissue trauma on resistance to local infection. An experimental study in rats. J Orthop Res 21:373-378.
33. Dan BJ, Kim SE, Pozzi A. 2014. Management of an infected cementless cup with prosthetic retention and antibiotic therapy in a dog. J Small Anim Pract 55:585-588.
34. Seibert RL, Lewis DD, Coomer AR, et al. 2011. Stabilisation of metacarpal or metatarsal fractures in three dogs, using circular external skeletal fixation. N Z Vet J 59:96-103.
35. . General Considerations for Animal Studies for Medical Devices Draft Guidance for Industry and Food and Drug Administration Staff.
36. Yuehuei H. An RJF. Animal Models in Orthopaedic Research: CRC Press
37. Arens D, Wilke M, Calabro L, et al. 2015. A rabbit humerus model of plating and nailing osteosynthesis with and without Staphylococcus aureus osteomyelitis. Eur Cell Mater 30:148-161; discussion 161-142.
38. Nau C, Seebach C, Trumm A, et al. 2016. Alteration of Masquelet's induced membrane characteristics by different kinds of antibiotic enriched bone cement in a critical size defect model in the rat's femur. Injury 47:325-334.
39. Rochford ETJ, Sabate Bresco M, Zeiter S, et al. 2016. Monitoring immune responses in a mouse model of fracture fixation with and without Staphylococcus aureus osteomyelitis. Bone 83:82-92.
40. Lairmore MD, Khanna C. 2014. Naturally occurring diseases in animals: contributions to translational medicine. ILAR J 55:1-3.
41. Gun G, Kues WA. 2014. Current progress of genetically engineered pig models for biomedical research. Biores Open Access 3:255-264.
42. Drummond GB, Paterson DJ, McGrath JC. 2010. ARRIVE: new guidelines for reporting animal research. Exp Physiol 95:841.
43. Bugnon P, Heimann M, Thallmair M. 2016. What the literature tells us about score sheet design. Lab Anim 50:414-417.
44. Lovati AB, Romano CL, Bottagisio M, et al. 2016. Modeling Staphylococcus epidermidis-Induced Non-Unions: Subclinical and Clinical Evidence in Rats. PLoS One 11:e0147447.
45. Eka A, Chen AF. 2015. Patient-related medical risk factors for periprosthetic joint infection of the hip and knee. Ann Transl Med 3:233.
46. Montgomery MK, Hallahan NL, Brown SH, et al. 2013. Mouse strain-dependent variation in obesity and glucose homeostasis in response to high-fat feeding. Diabetologia 56:1129-1139.
47. Park S, Rich J, Hanses F, et al. 2009. Defects in innate immunity predispose C57BL/6J-Leprdb/Leprdb mice to infection by Staphylococcus aureus. Infect Immun 77:1008-1014.

48. Ikejima S, Sasaki S, Sashinami H, et al. 2005. Impairment of host resistance to Listeria monocytogenes infection in liver of db/db and ob/ob mice. Diabetes 54:182-189.
49. Xiang L, Avila MM, Klemcke HG, et al. 2017. A novel rat model of orthopedic trauma for prehospital pain studies. The FASEB Journal 31:1087.1010-1087.1010.
50. La Fontaine J, Chen C, Hunt N, et al. 2016. Type 2 Diabetes and Metformin Influence on Fracture Healing in an Experimental Rat Model. J Foot Ankle Surg 55:955-960.
51. Prabhakara R, Harro JM, Leid JG, et al. 2011. Suppression of the inflammatory immune response prevents the development of chronic biofilm infection due to methicillin-resistant Staphylococcus aureus. Infect Immun 79:5010-5018.
52. Levi M, Keller TT, van Gorp E, et al. 2003. Infection and inflammation and the coagulation system. Cardiovasc Res 60:26-39.
53. Farnsworth CW, Schott EM, Jensen SE, et al. 2017. Adaptive Upregulation of Clumping Factor A (ClfA) by Staphylococcus aureus in the Obese, Type 2 Diabetic Host Mediates Increased Virulence. Infect Immun 85.
54. Moter A, Gobel UB. 2000. Fluorescence in situ hybridization (FISH) for direct visualization of microorganisms. J Microbiol Methods 41:85-112.
55. Varrone JJ, de Mesy Bentley KL, Bello-Irizarry SN, et al. 2014. Passive immunization with anti-glucosaminidase monoclonal antibodies protects mice from implant-associated osteomyelitis by mediating opsonophagocytosis of Staphylococcus aureus megaclusters. J Orthop Res 32:1389-1396.
56. Yokogawa N, Ishikawa M, Nishitani K, et al. 2018. Immunotherapy synergizes with debridement and antibiotic therapy in a murine 1-stage exchange model of MRSA implant-associated osteomyelitis. J Orthop Res 36:1590-1598.
57. Stadelmann VA, Potapova I, Camenisch K, et al. 2015. In Vivo MicroCT Monitoring of Osteomyelitis in a Rat Model. Biomed Res Int 2015:587857.
58. Pannucci CJ, Wilkins EG. 2010. Identifying and avoiding bias in research. Plast Reconstr Surg 126:619-625.
59. Aban IB, George B. 2015. Statistical considerations for preclinical studies. Exp Neurol 270:82-87.
60. Pearce GL, Frisbie DD. 2010. Statistical evaluation of biomedical studies. Osteoarthritis Cartilage 18 Suppl 3:S117-122.
61. Tande AJ, Patel R. 2014. Prosthetic joint infection. Clin Microbiol Rev 27:302-345.
62. Kapadia BH, Berg RA, Daley JA, et al. 2016. Periprosthetic joint infection. Lancet 387:386-394.
63. Meric G, Yahara K, Mageiros L, et al. 2014. A reference pan-genome approach to comparative bacterial genomics: identification of novel epidemiological markers in pathogenic Campylobacter. PLoS One 9:e92798.
64. Jefferson KK, Cramton SE, Gotz F, et al. 2003. Identification of a 5-nucleotide sequence that controls expression of the ica locus in Staphylococcus aureus and characterization of the DNA-binding properties of IcaR. Mol Microbiol 48:889-899.
65. Maira-Litran T, Kropec A, Abeygunawardana C, et al. 2002. Immunochemical properties of the staphylococcal poly-N-acetylglucosamine surface polysaccharide. Infect Immun 70:4433-4440.
66. Mack D, Fischer W, Krokotsch A, et al. 1996. The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear beta-1,6-linked glucosaminoglycan: purification and structural analysis. J Bacteriol 178:175-183.
67. Hudetz D, Ursic Hudetz S, Harris LG, et al. 2008. Weak effect of metal type and ica genes on staphylococcal infection of titanium and stainless steel implants. Clin Microbiol Infect 14:1135-1145.
68. Christensen GD, Simpson WA, Bisno AL, et al. 1982. Adherence of slime-producing strains of *Staphylococcus epidermidis* to smooth surfaces. Infect Immun 37:318-326.
69. Mack D, Rohde H, Harris LG, et al. 2006. Biofilm formation in medical device-related infection. Int J Artif Organs 29:343 -359.
70. Arrecubieta C, Lee M-H, Macey A, et al. 2007. SdrF, a *Staphylococcus epidermidis* Surface Protein, Binds Type I Collagen. J Biol Chem 282:18767-18776.
71. Bowden MG, Visai L, Longshaw CM, et al. 2002. Is the GehD Lipase from *Staphylococcus epidermidis* a Collagen Binding Adhesin? J Biol Chem 277:43017-43023.
72. Foster TJ, Geoghegan JA, Ganesh VK, et al. 2014. Adhesion, invasion and evasion: the many functions of the surface proteins of Staphylococcus aureus. Nat Rev Microbiol 12:49-62.

73. Hartford O, O'Brien L, Schofield K, et al. 2001. The Fbe (SdrG) protein of *Staphylococcus epidermidis* HB promotes bacterial adherence to fibrinogen. Microbiology 147:2545-2552.
74. Heilmann C, Thumm G, Chhatwal GS, et al. 2003. Identification and characterization of a novel autolysin (Aae) with adhesive properties from *Staphylococcus epidermidis*. Microbiology 149:2769-2778.
75. Cramton SE, Gerke C, Schnell NF, et al. 1999. The intercellular adhesion (*ica*) locus is present in *Staphylococcus aureus* and is required for biofilm formation. Infect Immun 67:5427-5433.
76. Heilmann C, Schweitzer O, Gerke C, et al. 1996. Molecular basis of intercellular adhesion in the biofilm-forming *Staphylococcus epidermidis*. Mol Microbiol 20:1083-1091.
77. Roche FM, Meehan M, Foster TJ. 2003. The *Staphylococcus aureus* surface protein SasG and its homologues promote bacterial adherence to human desquamated nasal epithelial cells. Microbiology 149:2759-2767.
78. Rohde H, Burdelski C, Bartscht K, et al. 2005. Induction of *Staphylococcus epidermidis* biofilm formation via proteolytic processing of the accumulation-associated protein by staphylococcal and host proteases. Mol Microbiol 55:1883-1895.
79. Christner M, Franke G, C. , Schommer N, N. , et al. 2010. The giant extracellular matrix-binding protein of *Staphylococcus epidermidis* mediates biofilm accumulation and attachment to fibronectin. Mol Microbiol 75:187-207.
80. Clarke SR, Harris LG, Richards RG, et al. 2002. Analysis of Ebh, a 1.1-megadalton cell wall-associated fibronectin-binding protein of *Staphylococcus aureus*. Infect Immun 70:6680-6687.
81. Williams RJ, Henderson B, Sharp LJ, et al. 2002. Identification of a fibronectin-binding protein from Staphylococcus epidermidis. Infect Immun 70:6805-6810.
82. Cucarella C, Solano C, Valle J, et al. 2001. Bap, a Staphylococcus aureus surface protein involved in biofilm formation. J Bacteriol 183:2888-2896.
83. Izano EA, Amarante MA, Kher WB, et al. 2008. Differential roles of poly-N-acetylglucosamine surface polysaccharide and extracellular DNA in *Staphylococcus aureus* and *Staphylococcus epidermidis* biofilms. Appl Environ Microbiol 74:470-476.
84. Qin Z, Ou Y, Yang L, et al. 2007. Role of autolysin-mediated DNA release in biofilm formation of *Staphylococcus epidermidis*. Microbiology 153:2083-2092.
85. Monecke S, Coombs G, Shore AC, et al. 2011. A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant Staphylococcus aureus. PLoS One 6:e17936.
86. Lourenco A, Coenye T, Goeres DM, et al. 2014. Minimum information about a biofilm experiment (MIABiE): standards for reporting experiments and data on sessile microbial communities living at interfaces. Pathog Dis 70:250-256.
87. Coenye T, Goeres D, Van Bambeke F, et al. 2018. Should standardized susceptibility testing for microbial biofilms be introduced in clinical practice? Clin Microbiol Infect 24:570-572.
88. Malone CL, Boles BR, Lauderdale KJ, et al. 2009. Fluorescent reporters for Staphylococcus aureus. J Microbiol Methods 77:251-260.
89. Yarwood JM, Bartels DJ, Volper EM, et al. 2004. Quorum sensing in Staphylococcus aureus biofilms. JBacteriol 186:1838-1850.
90. Kato F, Nakamura M, Sugai M. 2017. The development of fluorescent protein tracing vectors for multicolor imaging of clinically isolated Staphylococcus aureus. Sci Rep 7:2865.
91. Riool M, Dirks AJ, Jaspers V, et al. 2017. A chlorhexidine-releasing epoxy-based coating on titanium implants prevents Staphylococcus aureus experimental biomaterial-associated infection. Eur Cell Mater 33:143-157.
92. Zhang X, de Boer L, Heiliegers L, et al. 2018. Photochemical internalization enhances cytosolic release of antibiotic and increases its efficacy against staphylococcal infection. J Control Release 283:214-222.
93. Bateman BT, Donegan NP, Jarry TM, et al. 2001. Evaluation of a tetracycline-inducible promoter in Staphylococcus aureus in vitro and in vivo and its application in demonstrating the role of sigB in microcolony formation. Infect Immun 69:7851-7857.
94. Li D, Gromov K, Soballe K, et al. 2008. Quantitative mouse model of implant-associated osteomyelitis and the kinetics of microbial growth, osteolysis, and humoral immunity. J Orthop Res 26:96-105.

95. Nishitani K, Sutipornpalangkul W, de Mesy Bentley KL, et al. 2015. Quantifying the natural history of biofilm formation in vivo during the establishment of chronic implant-associated Staphylococcus aureus osteomyelitis in mice to identify critical pathogen and host factors. J Orthop Res 33:1311-1319.
96. Daghighi S, Sjollema J, Jaspers V, et al. 2012. Persistence of a bioluminescent Staphylococcus aureus strain on and around degradable and non-degradable surgical meshes in a murine model. Acta Biomater 8:3991-3996.
97. Oliver JD. 2005. The viable but nonculturable state in bacteria. J Microbiol 43 Spec No:93-100.
98. Ramamurthy T, Ghosh A, Pazhani GP, et al. 2014. Current Perspectives on Viable but Non-Culturable (VBNC) Pathogenic Bacteria. Front Public Health 2:103.
99. Proctor RA, von Eiff C, Kahl BC, et al. 2006. Small colony variants: a pathogenic form of bacteria that facilitates persistent and recurrent infections. Nat Rev Microbiol 4:295-305.
100. Zhai Z, Li H, Qin A, et al. 2014. Meta-analysis of sonication fluid samples from prosthetic components for diagnosis of infection after total joint arthroplasty. J Clin Microbiol 52:1730-1736.
101. Onsea J, Depypere M, Govaert G, et al. 2018. Accuracy of Tissue and Sonication Fluid Sampling for the Diagnosis of Fracture-Related Infection: A Systematic Review and Critical Appraisal. J Bone Jt Infect 3:173-181.
102. Drago L, Signori V, De Vecchi E, et al. 2013. Use of dithiothreitol to improve the diagnosis of prosthetic joint infections. J Orthop Res 31:1694-1699.
103. Xu Y, Larsen LH, Lorenzen J, et al. 2017. Microbiological diagnosis of device-related biofilm infections. APMIS 125:289-303.
104. Emerson JB, Adams RI, Roman CMB, et al. 2017. Schrodinger's microbes: Tools for distinguishing the living from the dead in microbial ecosystems. Microbiome 5:86.
105. Stoodley P, Conti SF, DeMeo PJ, et al. 2011. Characterization of a mixed MRSA/MRSE biofilm in an explanted total ankle arthroplasty. FEMS Immunol Med Microbiol 62:66-74.
106. Schoenrath F, Kikhney J, Kursawe L, et al. 2018. Life on the driveline: Molecular detection and fluorescence in situ hybridization-based visualization of microbial species in patients with left ventricular assist devices. J Heart Lung Transplant 37:163-166.
107. Mallmann C, Siemoneit S, Schmiedel D, et al. 2010. Fluorescence in situ hybridization to improve the diagnosis of endocarditis: a pilot study. Clin Microbiol Infect 16:767-773.
108. Santos Ferreira I, Kikhney J, Kursawe L, et al. 2018. Encapsulation in Polymeric Microparticles Improves Daptomycin Activity Against Mature Staphylococci Biofilms-a Thermal and Imaging Study. AAPS PharmSciTech 19:1625-1636.
109. Van de Vyver H, Bovenkamp PR, Hoerr V, et al. 2017. A Novel Mouse Model of Staphylococcus aureus Vascular Graft Infection: Noninvasive Imaging of Biofilm Development in Vivo. Am J Pathol 187:268-279.
110. Lubbert C, Wendt K, Feisthammel J, et al. 2016. Epidemiology and Resistance Patterns of Bacterial and Fungal Colonization of Biliary Plastic Stents: A Prospective Cohort Study. PLoS One 11:e0155479.
111. Pawar V, Komor U, Kasnitz N, et al. 2015. In Vivo Efficacy of Antimicrobials against Biofilm-Producing Pseudomonas aeruginosa. Antimicrob Agents Chemother 59:4974-4981.
112. Stewart PS, Zhang T, Xu R, et al. 2016. Reaction-diffusion theory explains hypoxia and heterogeneous growth within microbial biofilms associated with chronic infections. NPJ Biofilms Microbiomes 2:16012.
113. Heinrich A, Heyl KA, Klaile E, et al. 2016. Moraxella catarrhalis induces CEACAM3-Syk-CARD9-dependent activation of human granulocytes. Cell Microbiol 18:1570-1582.
114. Mouton JW, Theuretzbacher U, Craig WA, et al. 2008. Tissue concentrations: do we ever learn? J Antimicrob Chemother 61:235-237.
115. Parra-Sanchez A, Lugo J, Boothe DM, et al. 2006. Pharmacokinetics and pharmacodynamics of enrofloxacin and a low dose of amikacin administered via regional intravenous limb perfusion in standing horses. Am J Vet Res 67:1687-1695.
116. Tottrup M, Bue M, Koch J, et al. 2016. Effects of Implant-Associated Osteomyelitis on Cefuroxime Bone Pharmacokinetics: Assessment in a Porcine Model. J Bone Joint Surg Am 98:363-369.

**Figure legends**

**Figure 1. Visualization of bacterial microcolonies with fluorescence *in situ* hybridization (FISH) in a patient with infected non-union after open bone fracture; an overview of bone material (green autofluorescence) with adjacent blood and tissue material. B and C magnifications of the inserts in A and B, respectively, showing bacteria close to the bone (DAPI channel in blue, shown in C in black and white).**

**Figure 2. Schematic overview of the increasing complexity possible for testing antimicrobial strategies in preclinical *in vivo* models of ODRI. Increasing complexity of the model requires greater investment in expertise and instrumentation and may only be required as a technology reaches the late preclinical phase, or the intervention is particularly tied to a clinical situation in the target species.**

**Figure 3. FISH of a tissue section from a subcutaneous implant infection mouse model colonized with coagulase-negative staphylococci. (A) Overview shows the histology of the tissue stained by nucleic acid stain DAPI (blue) and background auto-fluorescence (yellow/green). (B) Higher magnification reveals differential colonization of the tissue with parts with only few bacterial cells and parts with biofilms, as detected by the pan-bacterial probe EUB338 (green). Note the difference in FISH-signal intensity in individual bacterial cells (B) corresponding to the ribosome content and therefore activity of the bacteria.**

**Figure 4. Abscess tissue from a mouse model infected with Enterococcus sp. (A) Overview shows infiltration by granulocytes (DAPI, blue) and bright auto-fluorescent material consistent with plant fibers (yellow/green). The insert (B) at higher magnification shows FISH of different bacterial populations colonizing the fibers. The identical microscopic field with separate microscopic channels reveals the strong autofluorescence of the fibers (B1, nonsense FISH probe NONEUB338, green), whereas bacteria of different morphologies are detected by the pan-bacterial FISH probe EUB338 (B2, magenta). Only one population is detected by the *Enterococcus*-specific FISH probe (B3, yellow). DAPI show some cell nuclei in addition to the bacteria (B4, blue). Although the mouse was only inoculated with *Enterococcus*, FISH revealed a multispecies infection possibly resulting from a gut perforation.**