

Reduction of Sample Size Requirements by Bilateral Versus Unilateral Research Designs in Animal Models for Cartilage Tissue Engineering

Patrick Orth, MD,^{1,2} David Zurakowski, PhD,³ Mauro Alini, PhD,^{4,5} Magali Cucchiari, PhD,^{1,5}
and Henning Madry, MD^{1,2,5}

Advanced tissue engineering approaches for articular cartilage repair in the knee joint rely on translational animal models. In these investigations, cartilage defects may be established either in one joint (unilateral design) or in both joints of the same animal (bilateral design). We hypothesized that a lower intraindividual variability following the bilateral strategy would reduce the number of required joints. Standardized osteochondral defects were created in the trochlear groove of 18 rabbits. In 12 animals, defects were produced unilaterally (unilateral design; $n = 12$ defects), while defects were created bilaterally in 6 animals (bilateral design; $n = 12$ defects). After 3 weeks, osteochondral repair was evaluated histologically applying an established grading system. Based on intra- and interindividual variabilities, required sample sizes for the detection of discrete differences in the histological score were determined for both study designs ($\alpha = 0.05$, $\beta = 0.20$). Coefficients of variation (%CV) of the total histological score values were 1.9-fold increased following the unilateral design when compared with the bilateral approach (26 versus 14%CV). The resulting numbers of joints needed to treat were always higher for the unilateral design, resulting in an up to 3.9-fold increase in the required number of experimental animals. This effect was most pronounced for the detection of small-effect sizes and estimating large standard deviations.

The data underline the possible benefit of bilateral study designs for the decrease of sample size requirements for certain investigations in articular cartilage research. These findings might also be transferred to other scoring systems, defect types, or translational animal models in the field of cartilage tissue engineering.

Introduction

IN ADVANCED TISSUE engineering approaches for articular cartilage repair in the knee joint, treatment effects are typically determined in translational animal models by comparing a novel technique either with untreated defects or with a standard procedure.¹⁻³ Cartilage defects of such control groups can be established either in contralateral joints of the same animal (bilateral design) or in ipsilateral joints of additional animals (unilateral design).⁴

Statistical power is a measure of the ability to distinguish between null and alternative hypotheses and increases with sample size, *ceteris paribus*.⁵ Bilateral study designs allow for a double increase in the number of treated limbs, although the quantity of experimental animals remains constant. However, due to the potential intraindividual dependence between fellow limbs due to common and

genetically determined host factors such as age, sex, weight, tissue characteristics, physical activity, or hormonal status,⁶ complex statistical analyses of such obtained data are necessary to eliminate this bias and possibly diminish the clinical relevance of the experimental investigation.^{4,6-8} This study design is not appropriate when both sides receive either an equal treatment (e.g., by systemic or intra-articular administration of a therapeutic agent) or different treatments that may reciprocally affect the opposite limb. On the other hand, independent unilateral setups require twice as many animals to obtain the same number of treated limbs. Besides, as statistical independence might additionally enhance the variability of interindividual comparisons,⁶ the required sample size may even be further increased. Therefore, from an ethic and economical standpoint, bilateral treatment might be superior to a unilateral strategy in certain studies. With regard to the experimental setup, researchers and

¹Center of Experimental Orthopaedics, Saarland University, Homburg/Saar, Germany.

²Department of Orthopaedic Surgery, Saarland University Medical Center, Homburg/Saar, Germany.

³Children's Hospital, Orthopaedic Surgery and Biostatistics, Harvard Medical School, Boston, Massachusetts.

⁴AO Research Institute, AO Foundation, Davos, Switzerland.

⁵Acute Cartilage Injury (ACI) Collaborative Research Program, AO Foundation, Davos, Switzerland.

methodologists are however limited by governmental or institutional constrictions, for example, the advice against bilateral interventions in experimental animals.⁹ As of yet, an investigation of required sample sizes for unilateral versus bilateral study designs is lacking in preclinical models for cartilage tissue engineering.

Thus, to better define such sample sizes, we performed a statistical analysis of the spontaneous repair of osteochondral lesions in the rabbit model, comparing the histological grading¹⁰ of the repair tissues for unilaterally and bilaterally operated animals. For both study designs, we determined the required sample sizes to detect distinct differences (i.e., effect sizes) in the histological score. These sample sizes were based on the intraindividual variability between sides and the interindividual variability between animals. Specifically, we hypothesized that a lower intraindividual variability following the bilateral strategy significantly reduces the required number of limbs. This, in turn, would result in less than half of animals needed to treat in a limited number of studies.

Materials and Methods

Study design

Standardized osteochondral defects ($n=24$) were created in the trochlear groove of 18 rabbits. In 12 animals, defects were created unilaterally with right and left joints alternating (unilateral design; $n=12$ defects), while defects were created bilaterally in 6 animals (bilateral design; $n=12$ defects). Animals were sacrificed after 3 weeks, and osteochondral repair was evaluated applying an established complex histological scoring system.¹⁰ Based on obtained intra- and interindividual variabilities, for a power of 80% and applying varying standard deviations, required sample sizes for the detection of a broad range of discrete differences in the histological total score value were determined.

Animal experiments

All animal procedures were approved by the local Governmental Animal Care Committee, and the principles of laboratory animal care were followed. Female Chinchilla bastard rabbits ($n=18$; Charles River, Sulzfeld, Germany) were kept in air-conditioned rooms with a constant temperature and a regular light/dark scheme. All surgeries were performed within the same period. The rabbits (mean BW: 3.0 ± 0.2 kg; mean age: 14 ± 2 weeks) were anesthetized by intramuscular injection of Rompun (0.2 mL/kg BW; Bayer, Leverkusen, Germany) and Ketavet (0.75 mg/kg BW; Pharmacia & Upjohn, Erlangen, Germany). The knee joint was entered using a medial parapatellar approach. The patella was dislocated laterally, and the knee flexed to 90° . Standardized cylindrical osteochondral defects ($n=24$; diameter 3.2 mm; depth 5.0 mm) were created in the femoral patellar groove with a manual cannulated burr with a flat tip (Synthes, Umkirch, Germany) and a custom-made drill stop. The defects were washed with phosphate-buffered saline; the patella was reduced; and the incisions were closed in layers. All animals were allowed full weight bearing postoperatively. No postoperative analgesia was administered. After 3 weeks, the rabbits were euthanized with pentobarbital (150 mg/kg BW; Merial, Hallbergmoos, Germany), and the

distal femurs ($n=24$) were dissected and fixed in 4% phosphate-buffered formalin.

Histological analyses

The distal femurs were trimmed and decalcified. Paraffin-embedded frontal sections ($5 \mu\text{m}$; $n=8-10$ per defect) of the defects were stained with safranin orange/fast green (safranin O; Fig. 1) according to the routine histological protocols.¹¹ All sections of the unilateral group and all sections of the bilateral group were stained using single batches of the staining solution. For the histological grading of the articular cartilage repair tissue, a total of 223 safranin O-stained sections were scored blinded by three observers using the standard scoring system described by Sellers *et al.*¹⁰ (Table 1), featuring a low intra- and interobserver variability.¹² The inverse score comprises nine individual parameters for single characteristics of cartilage repair, resulting in a total average point value ranging from 31 (empty defect and no repair response) to 0 points (complete regeneration).

Statistical analysis

The coefficient of variation (%CV) was calculated for each histological parameter to provide an estimate of variability in total histological scores for unilateral and bilateral study designs. Sample size requirements for detecting mean differences in total score points to evaluate the effects of osteochondral repair (i.e., effect size estimates) based on 80% power with pooled standard deviations of 2, 3, 4, and 5 points were determined using the Student *t*-test for studies in which one joint is operated per animal and the paired *t*-test for studies involving two joints per animal to account for within animal correlation (nQuery Advisor; Statistical Solutions, Solutions, Saugus, MA). Statistical analysis was performed using SPSS software (SPSS, Inc./IBM, Chicago, IL). Two-tailed values of $p < 0.05$ were considered statistically significant.

Results

Intra- and interindividual variability of osteochondral repair

Relative standard deviations of all individual parameters and of the total score were higher following the unilateral study design (Table 2). The resulting coefficients of variation expressed as percent (%CV) following the unilateral and bilateral study design are given in Table 3 to illustrate the differences in interindividual and intraindividual variabilities, respectively. When individual score parameters between both study designs were compared, all coefficients of variation were decreased when both joints per animal were operated (Fig. 1). This finding was most pronounced for matrix staining (7.1-fold), subchondral bone formation (4.5-fold), and architecture of the surface (2.3-fold). Of note, when only one joint per animal was operated, the %CV values for matrix staining (233.9%), defect fill (123.9%), and architecture of the defect (115.5%) exceeded 100%, indicating that standard deviations were higher than the mean values of the respective parameters (Table 3). Consequently, the variability of the total histological score value was 1.9-fold increased following the unilateral design when compared with the bilateral approach (26 versus 14%). Thus, total and

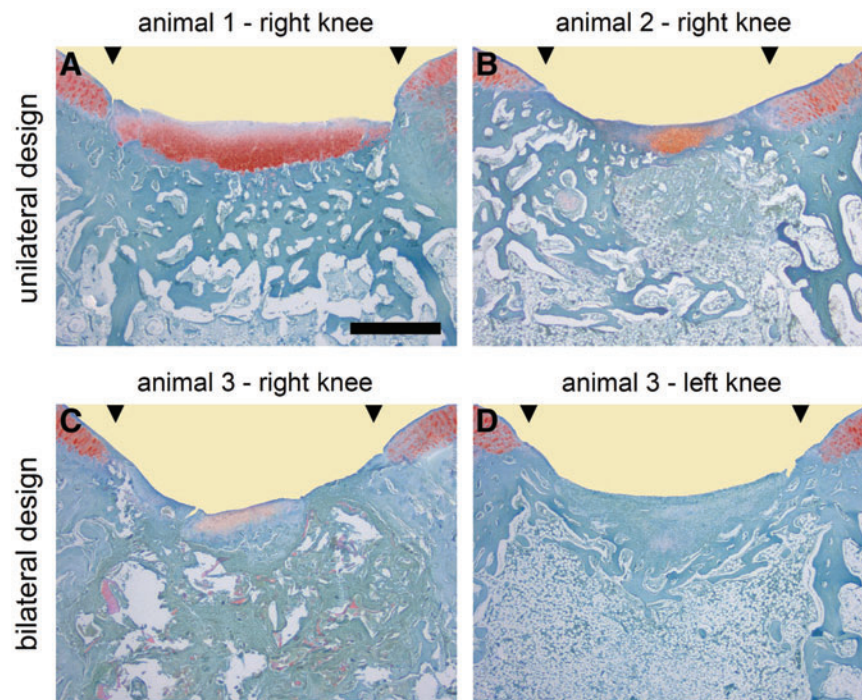


FIG. 1. Histological sections depicting interindividual and intraindividual variability of spontaneous osteochondral repair following a unilateral or bilateral study design. Representative safranin O-stained histological sections reflect the larger variability of spontaneous osteochondral repair for the unilateral experimental design compared with the bilateral design, resulting in increased sample size requirements when only one joint per animal is evaluated. In particular, the important histological parameters of matrix staining and subchondral bone reconstitution as well as the total score value¹⁰ vary more between different animals (unilateral design; images A and B) than between different joints of the same animal (bilateral design; images C and D; Table 3). Most single parameters of osteochondral repair, including matrix staining and subchondral bone reconstitution, as well as the overall histological score, were improved for the unilateral design compared with bilaterally operated animals (Table 2). The sections illustrated were taken from the defects having a histological rating equal to the mean score for its respective experimental group. Photomicrographs were obtained using standardized photographic parameters, including light intensity. Black triangles indicate defect margins. Scale bar: 1.0 mm. Color images available online at www.liebertpub.com/tec

individual score values are more homogeneously grouped when both joints of the experimental animal are analyzed (Fig. 1).

Sample size requirements for uni- and bilateral study designs

Assuming varying standard deviations between 2 and 5, sample size requirements to detect distinct effect sizes (differences in the total histological score value ranging from 1 to 10 points) were next determined for the unilateral (Table 4) and the bilateral study design (Table 5). Here, the numbers of required joints were always higher for the unilateral design (Table 4) compared with the bilateral strategy (Table 5). This effect was most pronounced when the aim was to detect small-effect sizes with large standard deviations (up to 2.0-fold increase in required limbs). In contrast, when large mean differences (between 6 and 10 points) are to be determined, the required sample sizes varied less between both study designs (1.0-fold to 1.5-fold increase in required joints).

For both experimental strategies, the largest sample sizes were required to detect the smallest-effect sizes: Assuming a standard deviation of three, for example, only three joints were necessary to detect a 10-point difference, while 143 joints were needed to detect a 1-point difference between

unilaterally created defects (48-fold increase; Table 4). In good agreement, the required sample size was 24-fold increased when comparing the detectability of smallest- versus largest-effect sizes following the bilateral design (3 versus 73 joints; Table 5).

Of note, Tables 4 and 5 display the required sample sizes in terms of joints or limbs. The differences in the sample sizes between the unilateral and bilateral study designs are further increased by factor 2 when transferring these results to the number of required experimental animals, featuring two treatable limbs per animal (Fig. 2).

Discussion

The data of the present study indicate that a bilateral study design, with one limb treated and the contralateral serving as a control, significantly decreases sample size requirements compared with a unilateral design (one treated limb per animal). Importantly, this effect is most pronounced when small-effect sizes are to be detected. Although these sample size requirements are based on different variabilities in the histological grading of lapine osteochondral repair, they might also be applicable to other scoring systems, defect types, or translational animal models in certain studies on cartilage tissue engineering.

TABLE 1. HISTOLOGICAL INVERSE GRADING SYSTEM FOR THE EVALUATION OF THE ARTICULAR CARTILAGE REPAIR TISSUES ACCORDING TO SELLERS *ET AL.*¹⁰

Filling of the defect relative to surface of normal adjacent cartilage	
111%–125%	1
91%–110%	0
76%–90%	1
51%–75%	2
26%–50%	3
<25%	4
Integration of repair tissue with surrounding articular cartilage	
Normal continuity and integration	0
Decreased cellularity	1
Gap or lack of continuity on one side	2
Gap or lack of continuity on two sides	3
Matrix staining with Safranin O-fast green	
Normal	0
Slightly reduced	1
Moderately reduced	2
Substantially reduced	3
None	4
Cellular morphology (chose first between a-b-c-d)	
(a) Normal	0
(b) Mostly round cells with the morphology of chondrocytes	
>75% of tissue with columns in radial zone	0
25%–75% of tissue with columns in radial zone	1
<25% of tissue with columns in radial zone (disorganized)	2
(c) 50% round cells with the morphology of chondrocytes	
>75% of tissue with columns in radial zone	2
25%–75% of tissue with columns in radial zone	3
<25% of tissue with columns in radial zone (disorganized)	4
(d) Mostly spindle-shaped (fibroblast-like) cells	5
Architecture within entire defect (not including margins)	
Normal	0
1–3 small voids	1
1–3 large voids	2
>3 large voids	3
Clefts or fibrillations	4
Architecture of surface	
Normal	0
Slight fibrillation or irregularity	1
Moderate fibrillation or irregularity	2
Severe fibrillation or disruption	3
Percentage of new subchondral bone	
90%–100%	0
75%–89%	1
50%–74%	2
25%–49%	3
<25%	4
Formation of tidemark	
Complete	0
75%–99%	1
50%–74%	2
25%–49%	3
<25%	4
Total score	0–31

TABLE 2. RESULTS OF THE HISTOLOGICAL ANALYSIS OF THE REPAIR TISSUE 3 WEEKS AFTER CREATION OF OSTEOCHONDRAL DEFECTS IN THE LAPINE TROCHLEAR GROOVE ACCORDING TO SELLERS *ET AL.*¹⁰

Parameter	Assignable points	Unilateral design	Bilateral design
Filling	[0–4]	0.476 ± 0.590	0.927 ± 0.821*
Integration	[0–3]	1.933 ± 0.800	1.229 ± 0.421*
Matrix staining	[0–4]	0.410 ± 0.958	2.802 ± 0.929*
Cellular morphology	[0–5]	2.295 ± 0.980	4.017 ± 0.975*
Architecture defect	[0–4]	0.800 ± 0.924	1.601 ± 1.074*
Architecture surface	[0–3]	1.333 ± 1.157	1.535 ± 0.598
Subchondral bone	[0–4]	1.124 ± 1.026	3.101 ± 0.636*
Tidemark	[0–4]	3.800 ± 0.526	3.903 ± 0.297
Total score	[0–31]	12.171 ± 3.215	19.115 ± 2.672*

Unilateral design: one joint operated per animal; bilateral design: both joints operated within the same animal. Values are given as mean ± standard deviation as well as range of point values (square brackets). * $p < 0.001$ for unilateral versus bilateral design.

The calculation of sample sizes is of paramount importance and mandatory before clinical or experimental investigations. Sample sizes are closely positively related to the statistical power of a study,¹³ in principle favoring the implementation of a large number of subjects. On the other hand, smaller sample sizes may translate into earlier completion of animal experiments or patient recruitment, thereby reducing the overall study costs and—finally—enabling patients to benefit from the conducted research sooner rather than later.¹⁴

TABLE 3. INTRA- AND INTERINDIVIDUAL DIFFERENCES IN THE HISTOLOGICAL GRADING OF THE LAPINE OSTEOCHONDRAL REPAIR TISSUE ACCORDING TO SELLERS *ET AL.*¹⁰

Parameter	Interindividual %CV (unilateral design)	Intraindividual %CV (bilateral design)	<i>x</i> -fold
Filling	123.88%	88.56%	1.40
Integration	41.37%	34.25%	1.21
Matrix staining	233.86%	33.16%	7.05
Cellular morphology	42.69%	24.27%	1.76
Architecture defect	115.51%	67.10%	1.72
Architecture surface	86.81%	38.39%	2.26
Subchondral bone	91.26%	20.51%	4.45
Tidemark	13.85%	7.60%	1.82
Total score	26.41%	13.98%	1.89

Interindividual coefficients of variation expressed as percent (%CV) were determined between animals ($n = 12$) with only one joint operated per rabbit. Intraindividual %CV was calculated between the right and left joints when both sides underwent surgical treatment (bilateral study design).

TABLE 4. SAMPLE SIZE REQUIREMENTS FOR A UNILATERAL STUDY DESIGN

Mean Difference	SD=5	SD=4	SD=3	SD=2
1 point	394	250	143	64
2 points	100	64	37	17
3 points	45	29	17	9
4 points	26	17	10	6
5 points	17	12	7	4
6 points	12	9	6	4
7 points	10	7	5	3
8 points	8	6	4	3
9 points	6	5	4	3
10 points	6	4	3	3

Required sample sizes when planning a study to assess differences in the total histological score values with one joint operated per animal (unilateral study design) are based on 80% statistical power using the two-sample Student *t*-test and display the numbers of joints needed to treat to detect the respective mean difference in the total histological score value according to Sellers *et al.*¹⁰ Mean differences are also applicable to any other scoring system, animal and defect model when one joint is operated per animal or treatment versus sham groups with one joint per animal. SD, standard deviation.

Here, the sensitivity of the applied scoring system could be increased from 10 points of detectable difference to 5 points by increasing the sample size from 3 to 7 joints in the bilateral design and from 3 to 6 joints in the unilateral design. Interestingly, to further improve the test sensitivity from 5 points to 1 point, required sample sizes increase to 73 for the bilateral and 143 limbs for the unilateral approach. Thus, if large mean differences of the obtained results are expectable, the analysis of unilaterally performed treatments requires a similar number of joints when compared with the bilateral approach. However, if one assumes minute-effect sizes, the bilateral design should be favored, as it results in an up to 2.0-fold reduction of the required limbs, thus decreasing the number of animals by factor 4.

TABLE 5. SAMPLE SIZE REQUIREMENTS FOR A BILATERAL STUDY DESIGN

Mean Difference	SD=5	SD=4	SD=3	SD=2
1 point	200	128	73	34
2 points	52	34	20	10
3 points	24	16	10	6
4 points	15	10	7	5
5 points	10	8	6	4
6 points	8	6	5	4
7 points	7	5	4	3
8 points	6	5	4	3
9 points	5	4	4	3
10 points	4	4	3	3

Required sample sizes when planning a study to assess differences in total histological score values with two joints operated per animal (bilateral study design) are based on 80% statistical power using the two-sample Student *t*-test and display the numbers of joints needed to treat to detect the respective mean difference in the total histological score value according to Sellers *et al.*¹⁰ Mean differences are also applicable to any other scoring system, animal and defect model when two joints from the same animal are used (one operated or treated and the contralateral as the control) to evaluate mean differences in score values between left and right sides.

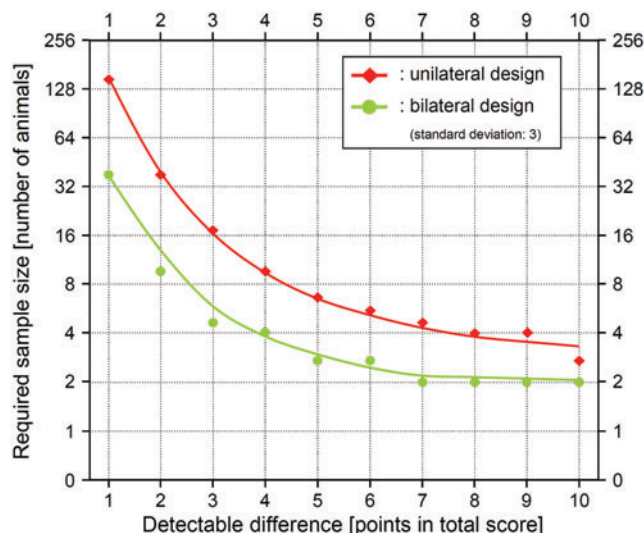


FIG. 2. Relation between the number of treated animals and the detectable differences in total point values. The number of animals needed to treat to detect specific differences in the total score value (effect size) is dependent on the study design: If only one joint per animal is treated (unilateral design), not only the double number of animals is required compared with a bilateral design (both joint operated), but also the animal sample size may have to be increased up to 3.9-fold (1-point difference). Here, a standard deviation of 3 is assumed, applying a histological score.¹⁰ Color images available online at www.liebertpub.com/tec

Medical research on the pairs of organs such as occurring in ophthalmology,^{8,13,15} nephrology,¹⁶ angiology,¹⁷ gynaecology,¹⁸ or orthopaedics^{7,19} is prone to the systematic error of multiple measurements in the same individual. As the statistical significance is more easily achieved with increased sample sizes, counting the dependent joints rather than independent animals or patients becomes attractive.^{4,19} However, due to common and genetically determined host factors such as age, sex, weight, tissue characteristics, physical activity, or hormonal status,⁶ these dependent observations are typically more similar than measurements in different, independent subjects. Therefore, the sample sizes for the bilateral study design presented here are not applicable when both sides receive an equal treatment (e.g., by systemic or intra-articular administration of a therapeutic agent), possibly diminishing the clinical relevance of this experimental setup. Instead, these sample size requirements may be applied to a limited number of experimental orthopedic investigations in which both sides are subjected to different treatments and in which the opposing (control) joint outcome is securely not affected by the therapy under investigation.⁴ Pathophysiological, molecular biological, or toxicological studies, as well as investigations on local tolerability, are also suitable for a bilateral design. Acknowledging the potential bias of dependency for intraindividual comparisons, statistical methods appropriate for data characterized by more than one source of variability (e.g., mixed models,⁶ generalized estimating equations,⁷ or marginal and conditional logistic regression models⁸) are required for a valid analysis in this bilateral setup. Yet, provided that the effect sizes are realistic and one can estimate the kind of variability (i.e., standard deviation) expected, the data on sample size

requirements for the unilateral or bilateral study design may be applied to any scoring system and animal model.

Apart from sample size, increasing the reliability of outcome measures can also enhance the statistical power.²⁰ Yet, it is more important for an outcome measure to be of practical importance²¹ and predictive for clinical applications²² than of the statistical difference. For example, evaluating exclusively the histological tidemark formation results in increased statistical power due to decreased variability of the data compared with the total score values. However, the resulting informative value would be limited. Besides, before any investigation, a realistic assessment of the minimum difference that would be worthwhile to detect^{13,23} is essential for the adaptability of the here-presented sample sizes. Another caveat to keep in mind when using these results is that sample size requirements were calculated assuming $\alpha=0.05$ and $\beta=0.20$ (i.e., power of 80%). If more stringent statistical requirements were needed, then either the sample size or the effect size would have to be increased.²³

Due to its simultaneous affection of both limbs, the bilateral study design is of limited suitability whenever postoperative unloading of the treated joint is necessary or when gait analyses are to be performed. Also, the animals employed here were in their late-juvenile stage, a factor that may possibly lead to improved cartilage repair in this population compared with older animals that may be more preferable for tissue engineering studies. Another limitation is that results obtained from the limbs of bilaterally operated animals cannot be compared conclusively with the limbs of unilaterally operated animals: although no differences in postoperative infection rates were reported after unilateral or bilateral tibial plateau leveling osteotomies in dogs,²⁴ the altered metabolic,²⁵ neurological,²⁶ pain-related,^{27,28} or biomechanical^{24,29} conditions rule out any valid comparability between these study designs. In the present study, particularly differences in the mechanical loading of the defects may have caused different osteochondral repair. Our data support this concept as the total score values differed significantly between unilaterally and bilaterally created defects. However, homogeneity rather than the magnitude of the score values was the focus here to determine sample size requirements.

In good agreement, significantly different outcomes have also been reported for bilateral versus unilateral procedures in clinical orthopedics. For example, bilateral total knee arthroplasty carries an increased risk of perioperative complications such as pulmonary fat embolism,³⁰ cardiac arrhythmia,³¹ and death¹⁴ compared to the unilateral procedures.³² These adverse effects outweigh the potential advantages of bilateral surgery such as significantly reduced overall costs of care, length of hospital stay, use of pain medication, and recovery time.^{33–35} Nevertheless, Bjorgul *et al.*³⁶ found that bilateral and unilateral cases of hip arthroplasty can be merged in heterogeneous sample groups of clinical trials without altering the findings compared with the homogenous study groups, justifying the analysis of bilateral procedures in the animal models for cartilage tissue engineering.

Conclusion

The data underline the benefit of bilateral research designs for certain studies on articular cartilage repair, demanding lower sample sizes than the unilateral design. Given the

premise that the different treatments do not reciprocally affect the opposite limbs in a bilateral experimental setup, these findings may be applied not only to cartilage repair analyses in tissue engineering, but also to other investigations on pairs of organs, possibly resulting in earlier completion of experiments and reduced study costs.

Acknowledgments

Financial support was received in part from the Deutsche Forschungsgemeinschaft (MA 2363/1–3 to HM), the Deutsche Arthrose-Hilfe e.V., and AO Exploratory Research of the AO Foundation. We thank E. Kabiljagic for help with the animal experiments.

Disclosure Statement

No competing financial interests exist for all authors.

References

1. Sniekers, Y.H., Weinans, H., Bierma-Zeinstra, S.M., van Leeuwen, J.P., and van Osch, G.J. Animal models for osteoarthritis: the effect of ovariectomy and estrogen treatment - a systematic approach. *Osteoarthritis Cartilage* **16**, 533, 2008.
2. Hunziker, E.B. Tissue engineering of bone and cartilage. From the preclinical model to the patient. *Novartis Found Symp* **249**, 70, 2003.
3. Sah, R.L., and Ratcliffe, A. Translational models for musculoskeletal tissue engineering and regenerative medicine. *Tissue Eng Part B Rev* **16**, 1, 2010.
4. Park, M.S., Kim, S.J., Chung, C.Y., Choi, I.H., Lee, S.H., and Lee, K.M. Statistical consideration for bilateral cases in orthopaedic research. *J Bone Joint Surg Am* **92**, 1732, 2010.
5. Solow, A.R., and Steele, J.H. On sample size, statistical power, and the detection of density dependence. *J Anim Ecol* **59**, 1073, 1990.
6. Ranstam, J. Problems in orthopedic research: dependent observations. *Acta Orthop Scand* **73**, 447, 2002.
7. Zhang, Y., Glynn, R.J., and Felson, D.T. Musculoskeletal disease research: should we analyze the joint or the person? *J Rheumatol* **23**, 1130, 1996.
8. Glynn, R.J., and Rosner, B. Comparison of alternative regression models for paired binary data. *Stat Med* **13**, 1023, 1994.
9. Auer, J.A., Goodship, A., Arnoczky, S., Pearce, S., Price, J., Claes, L., von Rechenberg, B., Hofmann-Amtinbrinck, M., Schneider, E., Muller-Terpitz, R., Thiele, F., Rippe, K.P., and Grainger, D.W. 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, 2007.
10. Sellers, R.S., Peluso, D., and Morris, E.A. The effect of recombinant human bone morphogenetic protein-2 (rhBMP-2) on the healing of full-thickness defects of articular cartilage. *J Bone Joint Surg Am* **79**, 1452, 1997.
11. Kiernan, J.A. *Histological and Histochemical Methods—Theory and Practice*. Oxford: Butterworth-Heinemann, 1999.
12. Orth, P., Zurakowski, D., Wincheringer, D., and Madry, H. Reliability, reproducibility, and validation of five major histological scoring systems for experimental articular cartilage repair in the rabbit model. *Tissue Eng Part C Methods* **18**, 329, 2012.
13. Goodall, E.A., Moore, J., and Moore, T. The estimation of approximate sample size requirements necessary for clinical

- and epidemiological studies in vision sciences. *Eye (Lond)* **23**, 1589, 2009.
14. Memtsoudis, S.G., Ma, Y., Gonzalez Della Valle, A., Mazumdar, M., Gaber-Baylis, L.K., MacKenzie, C.R., and Sculco, T.P. Perioperative outcomes after unilateral and bilateral total knee arthroplasty. *Anesthesiology* **111**, 1206, 2009.
 15. Leahey, A.B., Brucker, A.J., Wyszynski, R.E., and Shaman, P. Choriorretinal folds. A comparison of unilateral and bilateral cases. *Arch Ophthalmol* **111**, 357, 1993.
 16. Mao, Z., Ye, C., Mei, C., Zhao, X., Zhang, Y., Xu, C., Sun, L., Wu, J., Li, H., Dong, W., and Dong, S. Comparison of unilateral renal artery embolization versus bilateral for treatment of severe refractory hypertension in hemodialysis patients. *World J Urol* **27**, 679, 2009.
 17. Lee, A.Y., Gent, M., Julian, J.A., Bauer, K.A., Eriksson, B.I., Lassen, M.R., and Turpie, A.G. Bilateral vs. ipsilateral venography as the primary efficacy outcome measure in thromboprophylaxis clinical trials: a systematic review. *J Thromb Haemost* **2**, 1752, 2004.
 18. Morales-Ledesma, L., Betanzos-Garcia, R., and Dominguez-Casala, R. Unilateral or bilateral vagotomy performed on prepubertal rats at puberty onset of female rat deregulates ovarian function. *Arch Med Res* **35**, 279, 2004.
 19. Bryant, D., Havey, T.C., Roberts, R., and Guyatt, G. How many patients? How many limbs? Analysis of patients or limbs in the orthopaedic literature: a systematic review. *J Bone Joint Surg Am* **88**, 41, 2006.
 20. Leon, A.C., Marzuk, P.M., and Portera, L. More reliable outcome measures can reduce sample size requirements. *Arch Gen Psychiatry* **52**, 867, 1995.
 21. Gartland, J.J. Orthopaedic clinical research. Deficiencies in experimental design and determinations of outcome. *J Bone Joint Surg Am* **70**, 1357, 1988.
 22. de Vries, R.B., Buma, P., Leenaars, M., Ritskes-Hoitinga, M., and Gordijn, B. Reducing the number of laboratory animals used in tissue engineering research by restricting the variety of animal models. articular cartilage tissue engineering as a case study. *Tissue Eng Part B Rev* **18**, 427, 2012.
 23. Sumner, D.R., Turner, T.M., and Galante, J.O. Symmetry of the canine femur: implications for experimental sample size requirements. *J Orthop Res* **6**, 758, 1988.
 24. Fitzpatrick, N., and Solano, M.A. Predictive variables for complications after TPLO with stifle inspection by arthroscopy in 1000 consecutive dogs. *Vet Surg* **39**, 460, 2010.
 25. Karanth, S., and Nair, V.M. Changes in level of certain serum constituents following sciaticotomy in the common frog, *Rana tigrina* (Daud). *Indian J Exp Biol* **33**, 54, 1995.
 26. Wells, M.R., and Vaidya, U. Morphological alterations in dorsal root ganglion neurons after peripheral axon injury: association with changes in metabolism. *Exp Neurol* **104**, 32, 1989.
 27. Yasuda, T., Miki, S., Yoshinaga, N., and Senba, E. Effects of amitriptyline and gabapentin on bilateral hyperalgesia observed in an animal model of unilateral axotomy. *Pain* **115**, 161, 2005.
 28. Stevens, C.W., Kajander, K.C., Bennett, G.J., and Seybold, V.S. Bilateral and differential changes in spinal mu, delta and kappa opioid binding in rats with a painful, unilateral neuropathy. *Pain* **46**, 315, 1991.
 29. Cofone, M.A., Smith, G.K., Lenehan, T.M., and Newton, C.D. Unilateral and bilateral stifle arthrodesis in eight dogs. *Vet Surg* **21**, 299, 1992.
 30. Restrepo, C., Parvizi, J., Dietrich, T., and Einhorn, T.A. Safety of simultaneous bilateral total knee arthroplasty. A meta-analysis. *J Bone Joint Surg Am* **89**, 1220, 2007.
 31. Urban, M.K., Chisholm, M., and Wukovits, B. Are post-operative complications more common with single-stage bilateral (SBTKR) than with unilateral knee arthroplasty: guidelines for patients scheduled for SBTKR. *HSS J* **2**, 78, 2006.
 32. Stefansdottir, A., Lidgren, L., and Robertsson, O. Higher early mortality with simultaneous rather than staged bilateral TKAs: results from the Swedish Knee Arthroplasty Register. *Clin Orthop Relat Res* **466**, 3066, 2008.
 33. Kovacic, M.W., Singri, P., Khanna, S., and Gradisar, I.A. Medical and financial aspects of same-day bilateral total knee arthroplasties. *Biomed Sci Instrum* **33**, 429, 1997.
 34. Reuben, J.D., Meyers, S.J., Cox, D.D., Elliott, M., Watson, M., and Shim, S.D. Cost comparison between bilateral simultaneous, staged, and unilateral total joint arthroplasty. *J Arthroplasty* **13**, 172, 1998.
 35. Stubbs, G., Pryke, S.E., Tewari, S., Rogers, J., Crowe, B., Bridgfoot, L., and Smith, N. Safety and cost benefits of bilateral total knee replacement in an acute hospital. *ANZ J Surg* **75**, 739, 2005.
 36. Bjorgul, K., Novicoff, W.M., Brevig, K., Ahlund, O., Wiig, M., and Saleh, K.J. Patients with bilateral procedures can be included in total hip arthroplasty research without biasing results. *J Arthroplasty* **26**, 120, 2011.

Address correspondence to:

Henning Madry, MD
Center of Experimental Orthopaedics
Saarland University
Homburg/Saar D-66421
Germany

E-mail: henning.madry@uks.eu

Received: November 26, 2012

Accepted: March 14, 2013

Online Publication Date: April 25, 2013