
Kai Mithoefer¹, Daniel B. F. Saris², Jack Farr³, Elizaveta Kon⁴, Kenneth Zaslav⁵, Brian J. Cole⁶, Jonas Ranstam⁷, Jian Yao⁸, Matthew Shive⁹, David Levine¹⁰, Wilfried Dalemans¹¹, and Mats Brittberg¹²

Abstract

Objective: To summarize current clinical research practice and develop methodological standards for objective scientific evaluation of knee cartilage repair procedures and products. Design: A comprehensive literature review was performed of high-level original studies providing information relevant for the design of clinical studies on articular cartilage repair in the knee. Analysis of cartilage repair publications and synopses of ongoing trials were used to identify important criteria for the design, reporting, and interpretation of studies in this field. Results: Current literature reflects the methodological limitations of the scientific evidence available for articular cartilage repair. However, clinical trial databases of ongoing trials document a trend suggesting improved study designs and clinical evaluation methodology. Based on the current scientific information and standards of clinical care, detailed methodological recommendations were developed for the statistical study design, patient recruitment, control group considerations, study endpoint definition, documentation of results, use of validated patient-reported outcome instruments, and inclusion and exclusion criteria for the design and conduct of scientifically sound cartilage repair study protocols. A consensus statement among the International Cartilage Repair Society (ICRS) and contributing authors experienced in clinical trial design and implementation was achieved. Conclusions: High-quality clinical research methodology is critical for the optimal evaluation of current and new cartilage repair technologies. In addition to generally applicable principles for orthopedic study design, specific criteria and considerations apply to cartilage repair studies. Systematic application of these criteria and considerations can facilitate study designs that are scientifically rigorous, ethical, practical, and appropriate for the question(s) being addressed in any given cartilage repair research project.

Keywords

regulatory issues, general, clinical trial general, knee joint involved, cartilage repair repair, randomized controlled trials diagnostics

Introduction

Articular cartilage repair has seen an unprecedented clinical and scientific surge in the last decade due to an unmet clinical need facing a high incidence of articular cartilage lesions. Despite this renewed focus and the development of multiple new surgical techniques and cartilage repair products, the methodological quality of the existing clinical evidence for articular cartilage repair remains limited.¹² There is increasing recognition of the critical importance of the level of evidence for clinical study results that has led to efforts to improve the design of clinical studies evaluating established and novel cartilage repair technologies. This trend towards high-quality evidence-based science in articular cartilage repair is reflected by the publication of...
an increasing number of high-level studies. In general, properly designed and conducted randomized controlled clinical trial (RCT) designs produce the highest level of evidence upon which orthopedic surgeons can rely for their clinical decision making and therefore represent the recommended first choice in clinical study design. However, the relative paucity of isolated symptomatic chondral defects, frequent comorbidities, and resulting heterogeneous patient populations make conducting and completing evaluations of articular cartilage repair techniques and products very challenging. Patient enrollment and retention can be very difficult and requires screening large numbers of patients that can often only be achieved in a multicenter fashion. This difficulty is compounded by the fact that physicians and patients often lack equipoise when faced with a decision to enter into a randomized clinical study. Notwithstanding the methodological limitations found in the literature, current treatment algorithms provide a reasonable evidence basis for treatment decisions outside of randomized trials and may contribute to surgeon and patient reluctance to participate in randomized allocation of treatment in a trial.

The methodological difficulty of designing and conducting a rigorous and conclusive RCT is reflected by the fact that RCTs are rare in the orthopedic literature at large (11.3% of all studies), and those that exist often fail to meet all RCT study requirements. One meta-analysis of 2,468 randomized trials published in The Journal of Bone and Joint Surgery from 1988 to 2000 identified only 72 (2.9%) that met all the authors’ criteria for RCTs. A survey specific for the articular cartilage repair literature showed that only 6.6% of studies were RCTs, and many included significant methodological limitations such as small sample sizes. Development of methodology scores to evaluate the quality of clinical studies on articular cartilage repair has also been helpful in evaluating and improving the scientific quality of cartilage repair studies.

Since proper design of a cartilage repair trial is critical for valid and sound scientific conclusions, the International Cartilage Repair Society (ICRS) is aiming to support the development of appropriate clinical study designs. This report was prepared to outline the recommendations of the ICRS for conducting studies on articular cartilage repair based on the current literature, clinical practicability, and accepted standards of clinical care. The results of this review present a consensus statement of the ICRS and provide recommendations for the scientist conducting clinical research in articular cartilage repair in order to promote a level of standardization within the field. This review seeks to address the methodological limitations of the existing scientific evidence on articular cartilage repair based on the principles of randomized controlled studies and well-designed observational studies that have been found to be valuable data sources in the hierarchy of scientific evidence, particularly for surgical studies.

### Study Design

#### General Considerations

The primary purpose of controlled clinical trials on articular cartilage repair is to help assess the effectiveness of emerging technologies or surgical procedures to determine where they fall within the current standard of care for the treatment of chondral defects. Trials should be conducted in a way that minimizes risks while maximizing the benefits of the studied techniques for the patient and scientist. Common guidelines should apply to study design independent of whether the study is conducted by an independent clinical scientist or is industry sponsored. General guidance for cartilage repair product development has been provided by recent publications of the Food and Drug Administration (FDA) and European Medicines Agency (EMA) and is reflected in this ICRS consensus statement.

While regulations may differ by product designation, that is, medical devices, biologics, drugs, or combination products, similar principles hold for study designs evaluating them and should be aimed at developing and defining evidence-based treatment recommendations for articular cartilage repair. In general, study conduct must adhere to Good Clinical Practice (GCP) research standards and standards of clinical care and must adequately address the objective or purpose of the study while being feasible, ethical, and clinically relevant. Each type of study design has strengths and limitations, and no single study design will fit every product or surgical procedure and study objective.

Prospective RCT design is considered the highest level of clinical evidence. Aside from testing novel cartilage repair techniques or products, prospective RCTs can be used to test or reinforce findings from previous prospective and retrospective studies of established cartilage repair techniques with lower levels of evidence. While emphasis is placed on randomized trials as the optimal research design to evaluate a surgical intervention, RCTs are not always ethical or practical. Consequently, almost 90% of the orthopedic literature represents research findings from nonrandomized or observational study designs. Furthermore, despite the methodological advantages of an RCT, recent reports have identified nonsignificant differences between the results of nonrandomized and randomized designs. Recommendations for evidence-based study design should therefore not strictly be limited to evidence obtained from randomized trials only but should involve informed and effective use of all types of evidence. These principles of use of clinical evidence were applied as the basis for the current recommendations for clinical study design using the highest level of evidence available on each aspect of articular cartilage repair. While future studies should aim at conducting randomized controlled studies, the scientist designing a study on clinical articular cartilage repair needs to consider the practical and ethical aspects of surgical...
trials using GCP standards. Although an RCT is the most rigorous design, it may not provide all of the information a clinician might need.\textsuperscript{16} Observational studies may include a broader spectrum of patients (due to less rigid inclusion criteria), and therefore, the results may be more applicable to general orthopedic practice. Complementary use of observational studies and RCTs is therefore recommended. Each can provide useful information if the study adheres to careful study design with sound inclusion criteria for selecting patients. As recently pointed out, randomized controlled and observational studies together produce a more complete picture of the potential benefits and risks of a clinical decision for individual patients or health systems.\textsuperscript{16}

As in the design of any orthopedic clinical study, the initial study design planning for a cartilage repair trial should begin with a hierarchal assessment of the study objective(s) that addresses these key questions: What is the objective or hypothesis to test? What study designs can address this objective or hypothesis? Of these designs, which is the most scientifically rigorous? Is this design ethical and feasible? If not, what is the most rigorous design that is ethical and feasible? This assessment is designed to identify the most scientifically sound overall type of study design that is ethical, feasible, appropriate, and necessary to answer the hypothesis and meet the research objectives.\textsuperscript{23}

When RCTs are not feasible or not appropriate for the research objective, alternative controlled research designs should be considered. It is beyond the scope of this review to describe all possible designs in detail, but a brief description of recommended alternative designs and criteria is outlined here to provide researchers with an initial reference framework for clinical study planning. The ICRS recommendations for baseline data collection and outcome assessments are applicable to all of these study designs, although the level of detail may vary depending upon the scope and objectives of any particular trial. In addition to the RCT, the ICRS recommendations for appropriate trial designs for cartilage repair studies are cohort studies, registries, and within-patient control designs. In some circumstances, case control or historical control designs might apply, but since these are likely to be more unusual cases, these 2 designs are not addressed here in these general recommendations.

Cohort studies in the context of these ICRS recommendations are defined as comparative cohort studies. Treatment allocation is the key feature distinguishing a comparative cohort study from an RCT. In the cohort study, the treatment an individual receives reflects standard clinical practice at the study site(s) and is not chosen through a trial protocol random assignment process. Strengths of prospective comparative cohort designs include that it is easier to engage physician participation, patient enrollment is facilitated because patients are ensured that study participation does not otherwise affect the treatment they receive, and statistical methods such as regression analyses can be utilized to adjust for known confounders. Limitations of this design include that baseline data collection may not account for all potential confounding variables and even carefully performed analyses may not be able to adjust for these factors. Results of comparative cohort studies are more credible if the treatment effects are large.

Well-designed registries provide a valuable research tool for the longer term follow-up of broad patient populations undergoing cartilage repair procedures. The Agency for Healthcare Research and Quality (AHRQ) has published detailed guidance for prospective data collection and analysis plans to maximize the objectivity and value of registry designs.\textsuperscript{24} The Cartilage Repair Registry is an example of such a registry with a 15-year experience of collecting and publishing outcomes data consistent with these guidelines.\textsuperscript{25,26} Registries can be industry based, society based, and government based and may provide valuable datasets for research in articular cartilage repair. Whenever these registries are used, the origin of the datasets should be disclosed. Strengths of a registry include that it facilitates long-term data collection in a broad population that is more reflective of general orthopedic practice than other study designs. Uniform prospective data collection combined with appropriate prespecified analysis plans can reduce potential sources of bias that are encountered in less rigorous traditional observational case series. Limitations include that meeting these quality data collection objectives is expensive, privacy regulations and other administrative requirements are making it more challenging to sustain long-term registries and follow-up, and even well-designed and implemented analysis plans may not address all potential sources of confounding.

Within-patient control designs are also appropriate controlled study designs for some cartilage repair clinical research objectives. This type of design has been used in orthopedic research both in the context of total joint arthroplasty registries as well as stand-alone GCP trial design in cartilage repair.\textsuperscript{27} In a within-patient control design, the same patient receives 2 different treatments over time, and outcomes in the same patient are compared. It is particularly suited to assessing how a patient responds to a second-line or revision treatment after having failed a prior surgical treatment. Strengths of this design include that it minimizes variability due to patient factors and is an ethically and practical approach to assess a second-line treatment when randomization to a prior failed procedure is precluded. Limitations include that this design may have an inherent bias towards underestimating the efficacy and safety of the investigational treatment due to the challenges inherent in a revision environment and it tends to evaluate newer
technologies in the setting of older patients with more chronic conditions that may be suboptimal for cartilage repair.

Defining the treatment groups presents a critical step of designing a cartilage repair RCT and will impact the acceptability, applicability, and generalizability of the results by the public, scientific community, and regulatory agencies. Whether it is a single-, double-, or multiple-arm study, the treatment groups determine what inferences can be drawn from the results, the degree to which bias can be minimized, and the number and type of subjects that are recruited. Careful selection of the comparator (control) should permit quantifiable discrimination of subject cartilage repair outcomes resulting from the technique or product under investigation. For articular cartilage repair studies, a treatment arm using sham surgery is not recommended since failure to treat symptomatic cartilage defects conflicts with standards of clinical care for both medical and ethical reasons. Nonsurgical treatments (e.g., steroid injections, pharmaceuticals including nonsteroidal anti-inflammatories, nutraceuticals, viscosupplementation, or braces) typically are also considered inappropriate controls to compare cartilage repair procedures since most symptomatic subjects are likely nonresponders to these treatments prior to trial enrollment. Previous randomized and nonrandomized studies have utilized active comparators such as debridement/lavage, microfracture, or autologous chondrocyte transplantation.\(^5,8,11,12,25,28\) While specific outcomes related to these comparators are reported in the literature, arguably some of these studies have methodological limitations. However, more than a decade of decision making and surgeon experience with prospective follow-up is available for several articular cartilage repair options providing scientifically credible historic controls against which new technologies can be compared.\(^16,20\) Evaluation of these existing and new techniques in ongoing and future RCTs should provide more standardized results with a higher level of evidence, thus providing new perspectives on the existing scientific data.

**Primary and Secondary Study Endpoints**

Trial endpoints should be clinically relevant and address the overall objectives of the study. Primary endpoints should measure improvement in patient condition and may vary depending on the investigated product or technique. Pain and joint function present 2 critically important study endpoints and should be included in the study outcome assessment. The Knee injury and Osteoarthritis Outcome Score (KOOS) has been recently validated for articular cartilage repair with good psychometric properties.\(^29\) The KOOS subscores are recommended as basic assessment tools for the primary endpoints in studies on articular cartilage repair. The separate KOOS subscales of this patient-reported outcome instrument provide relevant and quantitative data including pain and joint function, the major components of primary outcome. Other useful measurement instruments that have been evaluated for articular cartilage repair are the International Knee Documentation Committee (IKDC) Subjective Knee Form, Western Ontario McMaster Universities Osteoarthritis Index (WOMAC), Modified Cincinnati Knee Rating System, Short Form 36, and the Lysholm scoring scale.\(^30-34\) The literature is lacking related to which specific primary outcome tool or subscale should be selected to assess patient condition, and clearly, more than one may be applicable and valid. Ideally, the scale chosen should also be utilized as the primary determinant for study inclusion at a predetermined pretreatment threshold of impairment. Finally, determining the meaningful clinical difference is also a subject of considerable debate, and exact parameters related to specific clinically relevant recommendations are lacking. Evaluation of postintervention activity levels compared to preintervention and preinjury levels may also provide clinically relevant information. Activity rating scales that have been used to provide quantitative analysis after cartilage repair procedures include the Tegner-Wallgren Activity Scale\(^35\) and Marx Activity Rating Scale\(^36\) and can be particularly helpful in the evaluation of athletic patients. It is recommended that at least one knee-specific and one generic measure be included in cartilage repair trials. Additional outcome instruments that can be used for the primary evaluation of articular cartilage repair are discussed in a more detailed, separate ICRS recommendation document on patient-reported outcome instruments in this volume.\(^37\) In addition to measurement of joint function and pain, macroscopic assessment of the repair cartilage has been validated for articular cartilage repair using the ICRS and Oswestry macroscopic cartilage evaluation scores\(^38-40\) (**Table 1**). These evaluations allow for graded quantitative analysis of the degree of defect repair, integration with the surrounding cartilage, and macroscopic appearance of the repair cartilage tissue. While macroscopic/microscopic evaluation may be recommended from a methodological standpoint, it is more difficult to apply in a trial design. For example, mandatory second-look arthroscopy with its associated surgical and anesthesia risks should be seriously considered from both clinical and ethical standpoints before entered into the protocol. Moreover, voluntary second-look arthroscopy or second-look arthroscopy in symptomatic patients may introduce selection bias and are therefore not recommended for routine primary endpoint evaluation.

Several qualitative and quantitative parameters that are clinically relevant to show difference in treatment groups but that are not validated for articular cartilage repair can be used for secondary endpoint assessment. Until there is a better understanding of the role of structural outcome...
parameters on long-term outcome and on how they relate to clinical outcome and other noninvasive methods of structural outcome evaluation, there can be an important role for including such evaluation tools in a research protocol. These parameters include clinical examination findings such as range of motion, degree of joint effusion, and locking or catching sensations. The assessment of structural properties of the repaired tissue by several methods can also serve as a secondary endpoint. First, articular cartilage repair tissue structure analysis using cartilage-specific magnetic resonance imaging (MRI) sequences can provide quantitative data on repair cartilage morphology and volume, peripheral integration, subchondral bone changes, and other structural characteristics. Marlovits et al. have proposed a magnetic resonance observation of cartilage repair tissue (MOCART) scoring system that sums up these types of quantitative MRI observations as a single overall score.\(^4^1\) MRI technologies such as delayed gadolinium enhanced magnetic resonance imaging of cartilage (dGEMRIC), T1rho, T2 mapping, magnetization transfer contrast, and diffusion weighted imaging can provide additional quantitative and qualitative information about the cartilage repair tissue.\(^4^2\) Advantages of MRI for secondary structural repair assessments are that it is noninvasive and can assess the entire repair area and surrounding tissue. MRI evaluation also allows for noninvasive longitudinal follow-up and comparison. The combination of simultaneous staged clinical endpoint determinations with arthroscopy for macroscopic assessment, microscopy, and advanced imaging technique will ultimately allow determination of which tools to select and apply in the “ultimate” study design.

Histological evaluation can also provide information about the quality of repair tissue and should follow the ICRS Histology Endpoint Committee guidelines. These include quantitative scoring systems (ICRS I and ICRS II) to evaluate tissue surface integrity, matrix organization and characterization, cell distribution and viability, calcified cartilage and subchondral bone morphology, and additional tissue characteristics\(^4^3\) (Table 2). Additional immunohistochemical evaluation for collagen type and noncollagenous matrix protein expression can provide qualitative information about the repair cartilage tissue. Tissue biopsies should be obtained at standardized, predefined time points after surgery to avoid the influence of repair cartilage tissue maturation on the histological parameters.\(^4^4\) These and additional histological and microscopic methods to evaluate biopsies obtained from cartilage repair tissue are discussed in detail in a companion ICRS recommendation paper in this volume.\(^4^5\) Limitations of histology evaluations include potential sampling biases resulting from inhomogeneity of

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Appearance</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.A</td>
<td>Level with surrounding cartilage</td>
<td>4</td>
</tr>
<tr>
<td>75% repair of defect</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>50% repair of defect</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>25% repair of defect</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0% repair of defect</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>I.B</td>
<td>100% survival if initially grafted surface</td>
<td>4</td>
</tr>
<tr>
<td>75% survival if initially grafted surface</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>50% survival if initially grafted surface</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>25% survival if initially grafted surface</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0% survival if initially grafted surface</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>II. Integration</td>
<td>Complete integration with surrounding cartilage</td>
<td>4</td>
</tr>
<tr>
<td>Demarcation border &lt;1 mm</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>75% integrated, 25% with notable border &gt;1 mm</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>50% integrated, 50% with notable border &gt;1 mm</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0%-25% integrated</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>III. Appearance</td>
<td>Intact smooth surface</td>
<td>4</td>
</tr>
<tr>
<td>Fibrillated surface</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Small, scattered fissures and cracks</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Small and large fissures</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Complete degeneration of graft area</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

Overall assessment and score:  
- Grade 1: Normal 12 points  
- Grade 2: Nearly normal 8-11 points  
- Grade 3: Abnormal 4-7 points  
- Grade 4: Severely abnormal 0-3 points
Statistical Considerations

Sample Size and Power. An important step in the planning and designing of a study is to calculate the number of patients to be recruited. The aim is to provide a patient number estimate needed for detecting a true treatment difference that limits the chance of a possible false claim (type I error), while also limiting the risk of not being able to detect this true difference (type II error), and at the same time allows for a feasible and economically possible trial. Related to these are also the associated ethical implications of a trial being underpowered (a high type II error) and the unnecessary risk to large numbers of patients, respectively. The limit set for the type I error (or significance level) is traditionally at \( \alpha = 5\% \), while the type II error rate (indicative for the power) is chosen to be between \( \beta = 10\% \) and \( \beta = 20\% \).

The sample size calculations are usually performed using the primary endpoint, which thus has to be defined for these calculations. Although such calculations are mostly based on endpoint-related and methodological assumptions, if correct, they support the rigor of the design and the clinical relevance and interpretation of the study results. They should consequently be clearly specified up front. The sample size estimate should be aimed at providing clinically significant results based on the appropriate statistical methods. However, even though the primary endpoint is of main concern when calculating statistical power, the foreseen power for other endpoints could also be examined in the light of the established sample size estimate.

Included in the considerations for limiting the overall probability of a false claim on the primary endpoint (type I error), as mentioned above, is multiplicity—the inflation of type I error due to either multiple testing of the same parameter (for instance, having an interim analysis) or due to multiple testing of several parameters of the same concept.\(^46\) These considerations should both be accounted for and adjusted for in the sample size calculation. For example, an interim analysis performed for advanced determination of efficacy creates multiplicity problems because this means that the endpoint is tested twice instead of once. The chance that at least one of these tests results in a false-positive outcome is greater than when the test is performed only once.

There are added considerations when performing sample size calculation for noninferiority trials (see below). The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) guideline\(^47\) recommends the use of half the conventional significance level for 1-sided tests. Additionally, for noninferiority designs, consideration needs to be given to the delta (percentage difference between investigational and control) to assess at what level the investigational results are “not worse than” the control. The sample size calculations are presented in detail both in the study protocol, the statistical analysis plan, when applicable, and the results report. Post hoc power calculations performed in connection with the interpretation of results should be avoided.\(^48\)

Multicenter versus Single-Center Design. Multicenter trials are often used for recruiting a sufficient number of subjects to satisfy a trial objective within a reasonable time frame.\(^47\) They are also used to provide a better basis for the generalization of its findings, with a broader range of clinical settings. The decision to conduct a multicenter study versus a single-center study depends on several key variables.\(^29\) For studies intended to support marketing approval of a cartilage repair product, multicenter studies help to ensure widespread applicability, safety, and efficacy of the product across regions, hospitals, surgeons, and subjects. Statistical considerations must be made with regards to site-by-site variation, and excessive variation in the number of subjects treated per site should be avoided. It should be described in the study protocol how the site-specific treatment effects will be estimated and tested in the statistical analysis. Heterogeneity in the outcomes by site complicates the interpretation and must be addressed statistically. Orthopedic studies that are not industry based must consider the feasibility of multicenter studies based on the resources required to design, initiate, and maintain the study conducted over time. The procedures followed in the study should

### Table 2. International Cartilage Repair Society II Visual Histological Assessment Scale

<table>
<thead>
<tr>
<th>Histological Parameter</th>
<th>Visual Analog Scale Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Tissue morphology (polarized light)</td>
<td>0-100</td>
</tr>
<tr>
<td>2. Matrix staining (metachromasia)</td>
<td>0-100</td>
</tr>
<tr>
<td>3. Cell morphology</td>
<td>0-100</td>
</tr>
<tr>
<td>4. Chondrocyte clustering</td>
<td>0-100</td>
</tr>
<tr>
<td>5. Surface architecture</td>
<td>0-100</td>
</tr>
<tr>
<td>6. Basal integration</td>
<td>0-100</td>
</tr>
<tr>
<td>7. Tidemark formation</td>
<td>0-100</td>
</tr>
<tr>
<td>8. Subchondral bone abnormalities/marrow fibrosis</td>
<td>0-100</td>
</tr>
<tr>
<td>9. Inflammation</td>
<td>0-100</td>
</tr>
<tr>
<td>10. Abnormal calcification/ossification</td>
<td>0-100</td>
</tr>
<tr>
<td>11. Vascularization in repair tissue</td>
<td>0-100</td>
</tr>
<tr>
<td>12. Surface/superficial assessment</td>
<td>0-100</td>
</tr>
<tr>
<td>13. Midzone/deep zone assessment</td>
<td>0-100</td>
</tr>
<tr>
<td>14. Overall assessment</td>
<td>0-100</td>
</tr>
</tbody>
</table>

Note: From Mainil-Varlet et al.\(^43\)
be standardized, and the aim should be to achieve the same distribution of subjects to treatment within each site. Single-site studies may be appropriate when a single orthopedic surgeon or group intends to conduct an institutional observational study. However, the strict study criteria and necessary number of patients to appropriately power an RCT in articular cartilage repair may be difficult to accomplish within a single institution.

**Superiority versus Noninferiority.** Most trials are designed to evaluate if one treatment is superior to another and should do so on the basis of a prespecified clinically appropriate primary endpoint and statistical analysis plan using hypothesis tests, with a prespecified significance level of \( \alpha \) per ICH guidelines. In some instances, a “noninferiority” trial may be an appropriate trial design and statistical approach. For example, it is an appropriate design if the chosen control arm has a reliably established quantitative treatment effect and the goal is to show that a new investigational treatment has similar efficacy to the control. In this trial design context, “noninferiority” is interpreted as “not worse than the comparator by more than a specified small amount.” This small amount is known as the noninferiority margin. The correct approach to show noninferiority is to prespecify the noninferiority margin in the study protocol.

A noninferiority trial can be designed to evaluate if the investigational treatment has better outcomes for prespecified secondary outcomes, even if the primary endpoint is designed to show that the investigational treatment is “non-inferior” than the control treatment for the primary endpoint. Such an approach may be especially attractive if the investigational treatment also has the potential to be better than the established comparator with regards to an additional important outcome measure such as structural repair parameters. In a trial in which the chosen comparator does not have a reliably established quantitative treatment effect, the trial should be a superiority design.

**Choice of Control Group.** A well-designed scientific cartilage repair study should use a control group and try to ensure that the control and the experimental groups are comparable with regard to factors that might have a significant effect on the outcome. In cartilage repair, these include age, body mass index (BMI), time since onset, defect age, gender, meniscal status, additional damage to the joint, previous surgery, location and size of the defect, leg alignment, and stability. Defining the treatment and control groups is a critical step of designing a cartilage repair trial with important implications for potential trial participants as well as the scientific community and regulatory agencies. Whether it is a single-, double-, or multiple-arm study, the treatment groups impact what inferences can be drawn from the results, the degree to which bias can be minimized, and the number and type of subjects that are recruited. Castro stated that placebo-controlled trials are justifiable when they are supported by sound methodological consideration and when their use does not expose research participants to excessive risk of harm. Placebo studies have been performed regarding surgical treatment of osteoarthritis but not in studies evaluating articular cartilage repair. For articular cartilage repair studies, a treatment arm using sham operation is generally not recommended since failure to treat symptomatic cartilage defects conflicts with GCP standards for both medical and ethical reasons. Nonsurgical treatments (e.g., steroid injections, pharmaceuticals including nonsteroidal anti-inflammatory medications, nutraceuticals, viscosupplementation, or braces) typically do not serve well as the appropriate control for cartilage repair procedures since many trial subjects were likely nonresponders to these treatments prior to trial enrollment. Evaluating the prognosis of asymptomatic early-stage cartilage lesions without a cartilage repair treatment may be one situation in which a sham operation or nonsurgical treatment could be considered appropriate controls. This is the clinical context for much of the current outcomes literature on 1-stage arthroscopic procedures such as microfracture in which lesions are relatively small, acute, and incidentally diagnosed at the time of other arthroscopic procedures such as anterior cruciate ligament (ACL) reconstruction or partial meniscectomy.

Pre-existing data from relevant, high-level published clinical studies may also be acceptable as controls. The use of these historical control groups have been principally supported by the EMA and can be advantageous for use in clinical trials by facilitating enrollment, when use of a placebo is difficult, or when problems are encountered for the willingness of patients to be randomized to one of the treatment arms.

There is not a single “gold standard” control applicable to all trials. There are several choices for appropriate controls based on many factors including treatment algorithms, indications, demographics of study population, published data, country, health care system, and specific surgeon/investigator/trial center experience. In general, cartilage repair trials use an active control and compare one surgical product or procedure to another. Potential active controls in cartilage repair trials include debridement, microfracture, autologous osteochondral grafts, or autologous chondrocyte implantation. Choice of a single active control simplifies statistical analyses and head-to-head comparisons but may pose enrollment challenges depending on how participating trial sites view indications for the 2 treatment arms.

Several current cartilage repair studies use microfracture as the control group. Microfracture is often considered a standard of care due to its frequent clinical use and relatively large clinical experience base compared to other techniques. Therefore, microfracture is considered an
appropriate control choice for many studies. However, it has not been rigorously tested in controlled trials, and the available level of evidence for this technique is still limited. Nonetheless, based on current literature and the reality of clinical practice in many regions, it is feasible to include microfracture as a cartilage treatment to which others are compared in studies evaluating treatment of chondral defects. As noted previously, there is no single “gold standard” comparator, and these recommendations should not be interpreted to mean that microfracture is the only choice for a control arm or that microfracture is always an appropriate choice. For example, in studies of deep osteochondral defects requiring bone grafting, microfracture is not routinely recommended as a control.55 While microfracture is not traditionally recommended for treatment of osteochondral defects, some consideration in this clinical setting may remain appropriate given the favorable outcomes described in recent clinical trials.5,8,11,12

**Randomization Process.** Randomization of subjects to treatment groups is done with the goal of eliminating or minimizing systematic differences between the groups as well as surgeon or patient selection bias. Both known and unknown confounding factors, such as age, BMI, or preoperative diagnosis, which may bias the analysis of outcomes if distributed systematically, should be randomly allocated. In practice, however, the procedure for performing the randomization can be complicated and, if not carried out properly, can result in incorrect subject allocation. For example, it is important that the randomization outcome is concealed until after patients consent to participate. Otherwise, systematic invitation of patients can generate systematic differences between the treatment groups. The purpose of randomization is to produce similarity, not equality, between treatment groups; often, that assumption is checked with hypotheses tests of the baseline. Randomization provides a sound basis for statistical inference, allowing generalization back to the common population from which the random groups were selected.

It should be recognized that confounding bias, a systematic underestimation, or overestimation of a treatment effect due to systematic imbalance of prognostic factors should not occur with randomization. Stratification can be used in the randomization procedure to assure that the allocation of patients is maintained evenly for important factors in the population. For example, randomization with stratification for lesion type implies one randomization list for each type. When randomization is used with stratification, resulting reports and analysis should reflect it.

**Blinding Procedures.** Subjective assessments and decisions can be affected by knowledge of treatment assignments.56 A single-blind study has blinded patients; a double-blind study has both blinded patients and investigators or sponsor staff. When feasible in surgical trials, blinding should be used to reduce assessment bias and improve outcome objectivity. For example, patients in the placebo or comparator group may expect less benefit from the treatment than patients in the experimental group, which could affect self-reporting of the outcome. Similarly, surgeons may be less likely to identify treatment responses in the placebo or comparator group. Such differences in management of patients, assessments, and interpretation of treatment responses are likely to bias the results. When it is clinically possible, blinding (i.e., unawareness of assigned treatment) should therefore be considered.

Surgical cartilage repair RCTs offer several challenges to blinding since the cartilage repair technique or product may have differences in invasiveness (arthroscopy v. mini-open v. arthrotomy), treatment process (2-stage interventions v. a single surgery), or postsurgical follow-up requirements such as variations in rehabilitation. When it is clinically feasible, study design should therefore aim to use comparative cartilage repair techniques with similar indications, invasiveness, treatment process, and equivalent postoperative rehabilitation protocols. This will facilitate maintenance of blinding throughout the course of the study. However, it may be impossible to blind patients and surgeons for a clinical trial based upon the nature of the compared technologies, for example, in comparisons of different techniques such as autologous chondrocyte implantation (ACI) to microfracture. In such instances, the design can include blinded outcome assessment by independent evaluators, such as blinded analysis of repair tissue biopsies or postoperative MRI to reduce assessment bias and optimize objectivity of the study results.8,12 Breaking of the blind should only be made when necessary for the patient’s care. Whenever this has been done, it should be documented, reported, and explained.

**Patient Recruitment**

Inadequate sample size is a common problem in cartilage repair studies. It is difficult to find patients who meet inclusion criteria, especially if these limit common comorbidities seen with full-thickness cartilage lesions. Patient recruitment is consequently the most difficult step when conducting clinical trials on articular cartilage repair. A survey of participating general practitioners revealed that forgetfulness and time pressures were the main factors inhibiting recruitment.57 The increasing need for documentation, controls, monitoring, and legal concerns carries with it many responsibilities that must be addressed prior to conducting the study in order to ensure that appropriate resources are available to fulfill study needs.

The rate for successful recruitment and inclusion can be increased in several ways. Most importantly, by increasing the general awareness of general practitioners and surgical
colleagues and patients that a cartilage repair trial is being conducted, potential patients can be identified and referred. To achieve this, prudent use of the Internet, society publications, public news, and word of mouth are strong media for mobilizing an enrollment effort. In addition, holding educational activities such as grand rounds or lectures in outlying geographic areas can raise awareness. Local physiotherapists aware of the study can also be helpful in identifying patients with chronic but focal knee issues. Setting up the outpatient clinic and logistics in such a way that rapid screening of relevant candidates can be performed effectively is critical. This will help to speed up inclusion through patient forums and study reputation. In research centers, it may be possible to establish procedures consistent with privacy regulations to screen hospital admissions, outpatient clinic lists, emergency room visits, and operative room schedules for patients who might fulfill the required criteria. Attention to patient confidentiality and Health Insurance Portability and Accountability Act (HIPAA) regulations remains critical during this process. Announcements in the medical and public press about the study opportunity utilizing print, radio, and television advertising in addition to physician contacts, patients’ public health records, and medical brochures can be beneficial. All of these forms of communication must adhere to regulatory guidelines locally and nationally. In addition, while these techniques may yield significant numbers of interested potential patients, it is best to include enough information to minimize the screening of large numbers of inappropriately self-selected or referred subjects.

Potential barriers to enrollment include the limitations set by the strict inclusion criteria of the study protocol such as defect size, defect location, concomitant pathology, or baseline pain scores. Other important limiting factors are the unwillingness of patients to accept the process of treatment group randomization or the considerable time, travel, and administrative requirements associated with the trial. Regulated trials face other challenges such as cost, the burden of documentation, and long follow-up periods that can induce trial fatigue among investigators and patients.

### Indications, Contraindications, and Special Considerations

#### Indications

The ideal patient for a clinical RCT on articular cartilage repair is a young patient with a symptomatic focal full-thickness chondral or osteochondral defect surrounded by normal cartilage in an otherwise healthy knee. However, this ideal lesion presents the exception rather than the rule. Observational studies may include a broader spectrum of patients (due to less rigid inclusion criteria), and therefore, the results may be more applicable to general orthopedic practice. Defect depth, location, and size remain critical variables considered when designing a clinical trial. Classification of the cartilage defect should be performed using a previously described ICRS system (Table 3). In addition, dividing the knee joint into sectors in both the frontal and lateral views makes it easier to describe lesion location, and a simplified mapping system has been agreed upon by ICRS. At present, ICRS grade 2 or less chondrosis is often not treated by cartilage repair but cartilage debridement and would not be considered a “repairable lesion” when considering unipolar and bipolar lesions. ICRS grade 3 or greater lesions are considered for treatment and defined as a “repairable lesion” (Table 3). Unipolar, by definition, means that one surface of the compartment has a treatable chondral lesion (ICRS grade 3 or greater) and the opposing joint surface has a lesion not requiring restoration (grade 2 or less) or no lesion at all.

Indication criteria for articular cartilage repair studies continue to evolve. Current indications provide a baseline for future indications that can develop based on evolving scientific evidence (Table 4). For example, first-generation...
autologous chondrocyte implantation (ACI) received FDA approval in 1997 for treatment of secondary, contained Outerbridge grade 3 and 4 defects of the femoral condyle and trochlea in patients >18 years of age. Since then, peer-reviewed studies have demonstrated the efficacy of this technique for expanded indications and include patients under the age of 18 years and over the age of 45 years; patients with large, bipolar, and patellofemoral defects; and patients requiring concomitant surgeries such as meniscus transplant.

Although Zaslav et al. indicated equal effectiveness of ACI as a secondary option after either chondral debridement or microfracture, Minas et al. showed poorer results of ACI after marrow stimulation. Several studies demonstrated that age of the defect, that is, time since onset of the cartilage injury of more than 12 to 36 months, had a clear negative effect on both MF and cellular therapy outcomes. The evidence-based evolution of cartilage repair procedure indications will also help to create scientifically based recommendations with treatment algorithms that can be used to design more targeted and effective future cartilage repair studies.

### Table 4. Indications, Contraindications, and Relative Considerations

#### Indications
Symptomatic focal full or near full-thickness (ICRS grades 3 and 4) chondral or osteochondral lesions of the femoral condyle, trochlea, and patella Primary or secondary arthritic cartilage repair

#### Contraindications
Advanced degenerative joint changes (joint space narrowing >50%)
Uncorrected axial malalignment >5° for femoral defects
Uncorrected patellar maltracking or instability for patellofemoral lesions
Uncorrected ligamentous instability
Age >60 years
Complete meniscal deficiency
Limited patient compliance
Tumor
Infection
Inflammatory arthropathy
Systemic cartilage disorder

#### Special Considerations
Defect size >2-4 cm² (for marrow stimulation, osteochondral autograft)
Body mass index >30
Partial meniscal deficiency (<50%)
Mild joint degeneration (joint space narrowing <50%)
Uncontained chondral lesions (depending on techniques used)
Recent medical treatment/surgery
Asymptomatic defects
Bipolar defects

### Table 5. Recommended Time Intervals before Study Inclusion

#### Nonoperative Treatment

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-articular injections:</td>
<td></td>
</tr>
<tr>
<td>Steroid injection</td>
<td>4 weeks</td>
</tr>
<tr>
<td>Viscosupplementation</td>
<td>3 months</td>
</tr>
<tr>
<td>Autologous blood-derived growth factor concentrate</td>
<td>3 months</td>
</tr>
<tr>
<td>Oral medications:</td>
<td></td>
</tr>
<tr>
<td>Anti-inflammatory medication</td>
<td>1 week</td>
</tr>
<tr>
<td>Glucosamine and chondroitin sulfate</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Systemic corticosteroid therapy</td>
<td>4 weeks</td>
</tr>
</tbody>
</table>

#### Operative Treatment

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial meniscectomy</td>
<td>6 months</td>
</tr>
<tr>
<td>Arthroscopic lavage</td>
<td>6 months</td>
</tr>
<tr>
<td>Chondroplasty (calcified cartilage intact)</td>
<td>6 months</td>
</tr>
<tr>
<td>Synovectomy</td>
<td>6 months</td>
</tr>
<tr>
<td>Lateral release</td>
<td>6 months</td>
</tr>
<tr>
<td>Meniscus repair</td>
<td>6 months</td>
</tr>
<tr>
<td>Ligament reconstruction (ACL, PCL, MCL, MPFL)</td>
<td>6 months</td>
</tr>
<tr>
<td>Tibial or femoral osteotomy</td>
<td>12 months</td>
</tr>
<tr>
<td>Microfracture</td>
<td>12 months</td>
</tr>
<tr>
<td>Chondroplasty (through calcified cartilage)</td>
<td>12 months</td>
</tr>
</tbody>
</table>

Note: ACL = anterior cruciate ligament; PCL = posterior cruciate ligament; MCL = medial collateral ligament; MPFL = medial patellofemoral ligament.

### Contraindications and Special Considerations

#### Inflammatory or Infectious Disease. Inflammatory joint diseases of the knee, either chronic or acute, present an absolute contraindication for involvement in articular cartilage repair studies. Active acute or chronic infection of the index knee also presents an absolute contraindication. Infection that has been completely cleared as indicated by the absence of clinical symptoms for 6 months, negative joint aspirations, and normal blood parameters may be considered for inclusion.

#### Meniscal Status. The consequences of meniscectomy are well documented in the literature with predictable cartilage degeneration over time with more progressive deterioration after lateral meniscectomy than in the medial compartment. Patients with a postmeniscectomized knee have an increased incidence of articular cartilage defects likely due to a compromised load-bearing state resulting in significantly increased tibiofemoral contact pressures. To date, no information exists on how meniscal status affects the outcome of articular cartilage repair. High-quality studies on articular cartilage repair should document any meniscal pathology and previous or concomitant treatment, such as meniscal transplant, meniscal repair, or meniscectomy. As articular cartilage injury is commonly associated with present or prior meniscal injury, a general guideline of acceptable partial resection is useful for exclusion/inclusion criteria for routine cartilage repair studies. Based on
laboratory data obtained from serial meniscal resection, partial resection of more than 50% or radial tears extending to the periphery have been shown to lead to changes of mean and peak contact forces equivalent to total meniscectomy.\textsuperscript{74,75} Meniscal resection of less than 50%, either in a previous procedure or concomitant with articular cartilage repair, would therefore be acceptable from a clinical and study design standpoint. During clinical studies, the amount of remaining intact meniscus at the time of articular cartilage repair treatment should be documented to allow for prospective evaluation of its effect on cartilage repair. Simultaneous meniscal repair represents standard clinical practice if a repairable meniscal tear is present and is clinically recommended to optimize contact forces and joint biomechanics. However, including patients into an RCT with simultaneous meniscal repair is not recommended since it introduces methodological conflicts due to its potential effect on outcome parameters, a more invasive nature than partial meniscectomy, and the variability of meniscal repair techniques and healing rates. Concomitant meniscal repair may be used as a separate nonrandomized arm in high-level RCT or observational studies to evaluate its effect on articular cartilage repair. However, the number of appropriately selected patients who have a concomitant repairable meniscus at the time of cartilage repair is very small. Some authors have demonstrated that combined meniscal allograft transplantation and cartilage restoration provides successful functional joint restoration for combined meniscal deficiency and articular cartilage injury.\textsuperscript{68,69,76} While this approach presents the recommended clinical approach for combined pathology, it cannot be used in RCTs since the effect of the meniscal transplantation cannot be separated from the effect of the articular cartilage repair. Comparison to control groups with either isolated meniscal transplantation or articular cartilage repair represents an appropriate study design for combined cartilage defects and meniscal deficiency.

**Ligamentous Stability.** Injuries to the articular cartilage surfaces frequently occur either acutely concomitantly with an initial trauma or result from chronic pathological joint mechanics in the ligament-deficient knee with associated recurrent instability episodes. Indeed, articular cartilage injuries have been described in 16% to 50% of acute injuries of the ACL.\textsuperscript{77-81} Acute injuries of the posterior cruciate ligament have been associated with articular cartilage injury in 52%.\textsuperscript{82} ACL deficiency has been shown to cause increased anterior translation with a posterior shift of cartilage contact biomechanics and resultant increase in shear forces, contact force concentration, and increased cartilage deformation.\textsuperscript{83-85} As a result, ACL deficiency increases the odds of developing an articular cartilage defect approximately 1% per month.\textsuperscript{86} ACL reconstruction significantly reduces increased tibial translation\textsuperscript{87} and has been shown to significantly reduce the incidence of articular cartilage and meniscal injury compared to ACL-deficient knees.\textsuperscript{88}

Given the demonstrated deleterious effects of joint instability and shear forces in particular on intact articular cartilage,\textsuperscript{89} uncorrected joint instability is included in the absolute exclusion criteria for studies investigating articular cartilage repair. Stability testing should be performed using the Lachman test and pivot-shift test, which have been shown to be most sensitive and specific for evaluation of ACL instability and correlate with functional outcome after ACL injury and reconstruction.\textsuperscript{90-92} Both tests should be performed under anesthesia and in comparison to the uninjured contralateral knee to obtain a more reliable measurement of joint stability. An asymmetric pivot-shift test result, unilateral increase of anterior tibial translation >5 mm (grade 2 Lachman test finding), or absence of an endpoint on Lachman testing indicate joint instability, and each presents criteria for exclusion.\textsuperscript{91,93} If available, instrumented testing with the KT-1000 arthrometer (MEDmetric, San Diego, CA) can be used as a supplement to asymmetric measurements of >5 mm indicating instability and resulting in exclusion. Patients with qualifying articular cartilage defects that have undergone surgical ligament stabilization more than 6 months prior and fulfill the above-mentioned stability criteria are also appropriate for study inclusion. While combined ACL and cartilage injury normally results in exclusion, subjects with concomitant ACL and articular cartilage injury may be included as a separate subcohort analysis or in studies addressing this specific population if the appropriate control groups are included.

**Alignment.** The detrimental effect of tibiofemoral malalignment for cartilage repair has been previously recognized with inferior results reported for osteochondral autograft transplantation and autologous chondrocyte transplantation in patients with varus malalignment.\textsuperscript{94,95} Laboratory data have demonstrated that malalignment leads to increased medial or lateral compartment joint contact stresses. In turn, increased contact stresses have been shown to negatively affect chondrocyte function, with decreased proteoglycan production and reduced type II collagen expression.\textsuperscript{96,97} Mechanical loading has also been found to affect chondrocyte function in native articular cartilage as well as in cell-based cartilage repair implants.\textsuperscript{98} While the negative effects of mechanical overload should be avoided for articular cartilage repair studies, mechanical stimulation is desirable to stimulate chondrocyte function and repair. Recent biomechanical studies demonstrated that the contact pressure is above the injurious pressure threshold for chondrocytes at varus alignment of more than 5°.\textsuperscript{99-101} Alignment should be assessed in all patients considered for study inclusion by measuring the mechanical and anatomical axis of the extremity in standing long-leg radiographs. Patients with malalignment >5° are eligible for inclusion.
also if they have undergone realignment procedures more than 12 months before inclusion into the study and meet the above-mentioned alignment criteria.

**Age.** A significant influence of age on the outcomes of articular cartilage repair procedures has been described for several cartilage repair techniques including microfracture, osteochondral autograft transplantation, autologous chondrocyte transplantation, characterized autologous chondrocyte transplantation, and matrix-assisted autologous chondrocyte transplantation. 28,55,102-104 Younger age resulted in better rates of return to sport participation with all surgical techniques. 105 Age-dependent qualitative and quantitative differences in metabolic activity, repair processes, and matrix synthesis in the repair cartilage are thought to be responsible for this effect.106,107 However, the reported age thresholds are inconsistent and vary between 25 to 45 years. 9,102-104,108-110 Based on the known effect of age, recent RCTs have included patients between 18 to 50 years. 5,8,11,12 However, recent prospective results have demonstrated significantly improved functional outcome in patients aged 45 to 60 years after cell-based articular cartilage repair with similar failure rates compared to younger patients. 60 Based on the current scientific evidence, age >45 years cannot be considered absolute exclusion criteria. While cell-based cartilage repair has been shown to be effective in patients <18 years, 61 inclusion of patients younger than 18 years in controlled trials is not routinely recommended due to the legal and practical implications related to the consent process and the ethical treatment of minors. Patient’s age between 18 to 60 years seems therefore appropriate for inclusion in cartilage repair studies. Subset analysis for age will facilitate generation of more objective data on the influence of age on articular cartilage repair and the possible definition of an upper threshold age for biological joint resurfacing.

**Diffuse Degenerative Joint Disease/Osteoarthritis.** Little information is available about how joint degeneration affects articular cartilage repair. While some studies demonstrated inferior results for articular cartilage repair of chronic cartilage defects, 61,109 other reports have shown that articular cartilage repair in patients with early osteoarthritis can provide lasting improvement of pain and joint function with histological and biochemical analyses showing no inhibition of the regenerative process by the degenerative joint environment. 111-113 However, the threshold level of joint degeneration at which restorative efforts become unsuccessful has not been established.

More data are needed on the effects of joint degeneration on articular cartilage repair. It is recommended that only patients with early signs of joint degeneration or none at all should be routinely included in prospective studies of cartilage repair and that patients with more advanced degenerative disease be considered for separate studies specifically designed to address this area. Radiographs routinely recommended for evaluation of joint degeneration should include weightbearing anteroposterior and lateral radiographs, Rosenberg flexion views, and Merchant views. Grading of joint degeneration has been traditionally done using radiographic grading scales, which are based on the radiographic extent of osteophyte formation and joint space narrowing. 114-116 Clinical validation studies have shown good correlation of radiographic grading with the degree of joint pain, biomarkers of joint degeneration, as well as a high sensitivity (91%-98%) for detection of advanced osteoarthritis.117,119 Thus, radiographic grading can be effectively used to exclude severe joint degeneration in cartilage repair studies. Joint space narrowing >50% is a radiographic sign of advanced joint degeneration used in all grading scales and presents a contraindication for inclusion into cartilage repair studies. In contrast, the radiographic presence of osteophytes is a nonspecific indicator of the severity of cartilage degeneration and is not recommended as an absolute exclusion factor. 120

While radiographic grading can be effectively used to exclude advanced osteoarthritis, only moderate correlation exists between early radiographic signs of joint degeneration and the degree of chondral injury seen at arthroscopy. 10,120 MRI is considered a sensitive tool for detection of chondral pathology and has been shown to be a valuable indicator for progression of joint degeneration and may be able to replace the routine radiographic screening in the future. 121-123

**Absolute Weight and BMI.** BMI, rather than absolute weight, is commonly used for evaluating the influence of weight on the surgical outcome. Increased body weight is an established risk factor for the development of osteoarthritis and cartilage degeneration 124-126 and is thought to result from increased cartilage matrix catabolism. 127 The effect of body weight on articular cartilage repair has not been thoroughly evaluated. Some studies have demonstrated that BMI >30 kg/m² is associated with worse functional outcome after microfracture, 128,129 suggesting that excessive body weight may negatively affect cartilage repair. However, whether BMI has a similar effect on other cartilage repair technologies is not clear from the available data. Until better data become available, patients with BMI >30 kg/m² can be included, but BMI data should be collected, and a separate subset analysis is recommended when feasible to clarify the influence of the BMI.

**Defect Characteristics**

**Lesion size.** Lesion size is an important variable to collect for data analysis and may be related to other important elements in study design such as symptoms, association with concurrent injuries, and enrollment screening. Smaller lesion size has been associated with better outcome after both microfracture and osteochondral autograft transfer,
while no association was found between defect size and autologous chondrocyte transplantation. Therefore, both microfracture and osteochondral autograft transplantation are recommended for use in relatively small defects (<2–4 cm²), while autologous chondrocyte transplantation or osteochondral allograft transplantation can be used for investigations involving both small and larger cartilage lesions. Importantly, lesion size should be measured after debridement of the lesion to stable cartilage margins intraoperatively. Estimation from preoperative imaging studies has not been shown to be reliable. Lesion type. The influence of the type of cartilage lesion on cartilage repair outcomes is not clear, with some reports showing significantly better outcome in chondral compared to osteochondral defects after both microfracture and osteochondral autograft transfer. In comparison, lesion type has not been shown to affect outcome after autologous chondrocyte transplantation. Based on the current knowledge about the individual techniques, chondral and osteochondral defects can be included in cartilage repair studies, but additional subanalysis may aid in interpretation of results. Similarly, consideration should be given to excluding uncontained defects in studies involving the first-generation microfracture technique.

Lesion location and number. Variable information is available from the literature on the effect of lesion location on articular cartilage repair. Some authors observed better results with lesions on the lateral femoral condyle after osteochondral autograft transfer. Other authors have found both better and worse results for defects of the medial femoral condyle after microfracture. No effect of defect location on outcome was seen after autologous chondrocyte transplantation. Isolated grade 3 or 4 defects involving a single compartment of the knee represent optimal cases in cartilage repair studies. Multiple defects within a single compartment that are unipolar and involve consistent joint surfaces (femur, trochlea, and patella) represent the next level of complexity. Combined defects of different femoral condyles and combination of cartilage lesions of the femoral condyle and patellofemoral compartment represent complex surgical challenges and analytical issues for surgical trials. These types of cases are currently best addressed in limited studies in highly specialized centers and are not yet appropriate for routine broadly based trials. Due to the limited scientific evidence and different biomechanical and morphological considerations for tibial articular cartilage and lesions on opposing surfaces (“bipolar lesions”), these types of indications are not part of routine clinical practice or broadly based cartilage repair trials. However, studies using well-designed study arms for bipolar or isolated tibial lesions are encouraged to address the current lack of scientific evidence for these challenging cartilage lesions.

Concomitant Procedures. The importance of addressing concomitant pathology such as joint instability or malalignment for the outcome after articular cartilage repair has been well documented. Concomitant pathology is normally addressed in a simultaneous or staged approach. Simultaneously addressing all combined pathologies limits the amount of postoperative rehabilitation and time until functional restoration is complete. Concomitant procedures such as ACL reconstruction or osteotomy did not negatively affect outcome after autologous chondrocyte transplantation, while simultaneous adjuvant procedures were associated with better results after osteochondral autograft transfer and microfracture. However, concomitant procedures may introduce significant confounding factors that need to be addressed with methods such as excluding these concurrent procedures from high-level studies, adequate sample size and analytical plans to evaluate the potential effect of these simultaneous procedures on the primary treatment of interest, or staging these procedures with adequate time interval. If sample size, time, and resources permit, separate arms may be included into these studies that analyze the effect of individual well-defined simultaneous or staged surgical procedures.

Preoperative Activity Level. Higher preoperative activity levels have been associated with better clinical outcomes in randomized prospective studies, irrespective of the technique used. However, based on the available data, no single specific minimum or maximum functional level can be recommended as an absolute parameter for exclusion or inclusion. Pretreatment activity, function, and pain levels can be assessed using patient-reported outcome instruments recommended for primary outcome evaluation such as KOOS subscales, IKDC Subjective Knee Form, WOMAC, Modified Cincinnati Knee Rating System, Short Form 36, Lysholm scoring scale, Tegner-Wallgren activity scale, and Marx-Activity rating scale. Preoperative function and activity levels for inclusion should be considered and defined for each study depending on the individual study design and goal.

Preoperative Symptoms: Level and Duration. No information is specifically available from the literature on the effect of preoperative symptom severity on postoperative results from articular cartilage repair. Recent studies have shown that asymptomatic articular cartilage defects may do well without surgical intervention. Therefore, it is recommended that asymptomatic patients be excluded from cartilage repair studies. While no single degree of symptoms is specifically recommended for inclusion, a minimum severity of symptoms should be defined before the start of the study using quantitative evaluation of symptoms such as pain and/or effusion. Asymptomatic patients may be considered for inclusion as a separate control cohort. Previous studies have indicated that the preoperative duration of symptoms can affect functional outcome and postoperative activity.
levels after cartilage repair.\textsuperscript{61,103,105} No recommendation can currently be made on the maximum or minimum duration of symptoms to be included based on the existing literature. However, the duration of symptoms should be determined and recorded as part of the study design.

Contralateral and Ipsilateral Joint Pathology. The status of the joint in the ipsilateral and contralateral limb should be taken into consideration as part of the inclusion criteria. Both the ipsilateral and contralateral hip and ankle joints as well as the contralateral knee should be asymptomatic or minimally symptomatic, particularly for studies with pain as the primary endpoint.\textsuperscript{141} That is, the patient is not receiving treatment for any pathological conditions of any of these other joints. Exceptions include conditions that may typically occur bilaterally, such as patellofemoral dysplasia with associated chondrosis. These cases may be acceptable for inclusion with adequate subset design and analysis or when these findings are not clinically active or contributing to pain or impairment.

Prior Treatment. Prior treatment of the affected joint may affect the symptoms and outcomes for cartilage repair studies. To minimize the effect of prior treatments, the following timelines for inclusion are recommended based on empirical and objective criteria (Table 5).

Nonoperative treatment: glucocorticoid injection. Based on mean elimination half times reported for pharmacological studies, triamcinolone,\textsuperscript{142} betamethasone,\textsuperscript{143} and methylprednisolone,\textsuperscript{144} patients can be included into cartilage repair studies if they did not receive glucocorticoid injections within 4 weeks prior to enrollment.\textsuperscript{145}

Nonoperative treatment: intra-articular viscosupplementation. Based on mean elimination kinetics reported for the human knee in pharmacological studies, patients who received hyaluronic acid injections within 3 months of the study treatment should be excluded.\textsuperscript{146}

Nonoperative treatment: autologous blood-derived growth factor concentrate. Concentrates of autologous blood growth factors are increasingly used for treatment of musculoskeletal and degenerative joint conditions.\textsuperscript{147-149} However, limited systematic information is available on the exact growth factor concentrations in these injections, the elimination half-life in the joint, their local and systemic anabolic effects, the influence on intra-articular cytokines, and their inhibitory or synergistic effects on articular cartilage repair.\textsuperscript{150-154} In addition, the variable content of cellular components (i.e., leukocytes) in the different commercially available autologous blood concentrates introduces additional confounding factors. Based on the currently available scientific and empirical information, it is recommended that patients be excluded from articular cartilage repair studies if they have received these injections within 3 months of their study treatment. The prior use of these injections should also be recorded during the study.

Nonoperative treatment: oral medications

- Glucosamine and chondroitin sulfate: Previous reports have described the bioavailability and pharmacokinetics of glucosamine\textsuperscript{155} and chondroitin sulfate.\textsuperscript{156} Based on these data, glucosamine and chondroitin sulfate should be discontinued at least 2 weeks prior to treatment in any articular cartilage repair study.

- Oral anti-inflammatory medication: Limited and conflicting information is available on the effect of anti-inflammatory medication on proteoglycan synthesis and chondrocyte proliferation \textit{in vitro}, and the effect of these medications on articular cartilage repair \textit{in vivo} is not known.\textsuperscript{157-160} Based on the published pharmacokinetics of ibuprofen,\textsuperscript{161} naproxen,\textsuperscript{162} nabumetone,\textsuperscript{163} diclofenac,\textsuperscript{164} celecoxib,\textsuperscript{165} capsaicin,\textsuperscript{166} and ketorolac,\textsuperscript{167} patients should have discontinued these medications at least 1 week before their treatment. These medications should also be discontinued 1 week before follow-up evaluations to avoid potential interference with measurement primary and secondary endpoint parameters.

- Systemic corticosteroid therapy: Due to the pharmacology of systemically applied glucocorticoids,\textsuperscript{168} patients who have received either hydrocortisone, prednisolone, methylprednisolone, or dexamethasone within 4 weeks of treatment in a study should be excluded.

Operative treatment. The influence of prior surgical treatment on subsequent cartilage repair is still controversial. While some authors have reported an influence of the prior surgery on the outcome of the secondary procedure,\textsuperscript{61,70} others have not shown any effect.\textsuperscript{59} The potential for selection bias can be minimized by following appropriate randomization procedures. Detailed documentation of the prior procedures and subset analysis is recommended. Prior procedures may present technique-specific contraindications for cartilage repair studies. Patients with prior cartilage repair procedures that remove the subchondral bone plate (i.e., osteochondral grafting) should not participate in subsequent trials evaluating cartilage repair techniques that require an intact subchondral bone plate (i.e., microfracture). Appropriate time intervals from the prior surgeries (i.e., adjuvant procedures, prior cartilage repair procedures, or unrelated surgeries) are required in order to avoid their influence on the baseline evaluation of the cartilage repair study. Based on the empirical time required for postoperative recovery, the following minimum intervals are recommended before inclusion into articular cartilage repair studies.
Special Considerations for Study Design in Articular Cartilage Restoration in the Patellofemoral Compartment

Some unique features of the patellofemoral (PF) compartment have implications for study design for articular cartilage repair in this compartment. PF pain presents a very complex clinical picture and is often modulated through and combined with sources of pain other than the pathological cartilage tissue. Due to the complex nature of PF cartilage defects, it is necessary to identify and address all other sources of pain and design a comprehensive treatment and study design plan for each problem. However, this introduces a variable combination of confounding factors into studies investigating PF cartilage restoration. While it would be desirable to eliminate all confounding factors when designing a PF cartilage study, this proves to be more difficult than for the tibiofemoral (TF) compartment. Furthermore, a larger percentage of the PF patient pool may have bilateral pathology when compared to TF patients. This makes the use of standard outcomes tools difficult, as most of the tools assume that one limb is normal. One approach is to study only those patients who have unilateral pathology, yet that would eliminate many clinically relevant scenarios, such as static chronic patellar lateral positioning (chronic patellar subluxation) or the sequela of patellar instability, and increase screening and enrollment challenges. Therefore, PF cartilage restoration studies will typically require a much larger pool of patients to adequately power subset analysis of the subgroups.

The peer-reviewed literature demonstrates the safety and efficacy of current PF cartilage restoration techniques.\textsuperscript{169-171} Review of the applications of these techniques will be useful when contemplating analogous new cartilage restorations for a particular PF cartilage pathology. To optimize cartilage restoration outcomes in the PF compartment, it is essential that all other pathologies in the compartment are considered and treated to optimize the environment for the cartilage implant. The factors to be considered include the medial soft tissues, lateral soft tissues, and tibial tuberosity position.

**Medial soft tissue considerations.** Chondral injuries in the PF compartment can result from acute or chronic patellar instability. Lateral patellar dislocation results in injury not only to the medial soft tissue restraints but is often associated with PF chondral injury.\textsuperscript{172} As the chondral injury is typically distal medial, it is important to normalize the medial soft tissue restraints without overloading the medial PF compartment. Nonanatomical medial repairs have resulted in a high incidence of arthrosis.\textsuperscript{173,174} The primary restraint to lateral displacement forces is the medial patellofemoral ligament (MPFL) with lesser contributions coming from the medial patellomeniscal ligament (MPML) and medial patellotibial ligament (MPTL). Whether the MPFL is repaired, shortened, or reconstructed, the goals of medial soft tissue surgery are prevention of lateral instability and avoiding medial overload.

**Lateral soft tissue considerations.** The static lateral patellar restraints are the 2 layers of the lateral retinaculum. The indication for lateral release (LR) is clinical symptomatic patellar tilt, which has been documented by MRI or computed tomography. The decision for an LR should be based on the status of the MPFL, degree patellar tilt, tibial tuberosity position, location of the cartilage defect, and location of pain. Routine or overly extensive LR should be avoided since it may not improve PF contact stresses and can lead to iatrogenic medial instability.\textsuperscript{175,176}

**Tibial tuberosity considerations.** The tibial tuberosity plays a key role in alignment and patellar tracking. Excessive lateral position of the tibial tuberosity relative to the trochlear groove (TT-TG distance >15-20 mm) results in increased lateral force vectors and contact forces on the patella cartilage during knee motion.\textsuperscript{177} The goal is to optimize force and contact area through repositioning the tuberosity medially, laterally, proximally, and/or distally depending on the individual cartilage defect location and tibial tubercle location.\textsuperscript{178-180} A precise preoperative plan should also include determination of the vertical patellar height using the Caton-Deschamps index or Blackburn-Peel index. With each tuberosity surgery, the goals are to normalize anatomy and optimize stress. The importance of optimizing PF stress has been demonstrated by early reports of cell-based PF cartilage restoration showing poor results without stress optimization, while later studies using the same cartilage restoration technique but with attention to decreasing PF stress demonstrated much improved clinical results.\textsuperscript{179-182}

Given the complexity of PF compartment presentations and that current clinical standards to optimize PF compartment cartilage restoration outcomes include soft tissue and osteotomy procedures to correct patella alignment and stability, it is understandable that most PF cartilage repair studies to date have reported results with the use of the concomitant procedures.\textsuperscript{179-185} In order to better understand the contribution of patella resurfacing procedures to the overall outcomes of complex PF surgeries, the design of future PF studies should consider more detailed documentation of the indications for and quantitative results of the concurrent alignment procedures. For example, medial restraints should be measured in quadrants of lateral displacement, lateral retinaculum...
procedures should be based on imaging proven tilt, and tibial tuberosity osteotomies should include TT-TG measurements. This will improve comparability of future studies and allow for a more evidence-based objective evaluation of the confounding factors and cartilage repair technologies in the complex clinical setting of articular cartilage injury in the PF compartment. Despite the general limitations of historical controls, the assessment of patella cartilage resurfacing with contemporary standards of concurrent alignment procedures may be a specific case in which a historical control of alignment without resurfacing presents an appropriate study design until improved data are available.

Summary and Conclusion

New biological approaches have stimulated extensive interest and developments in articular cartilage repair. Despite this renewed focus with multiple new surgical techniques and cartilage repair products evolving, the methodological quality of the existing clinical evidence for articular cartilage repair remains limited. High-quality clinical research methodology is critical for the optimal evaluation and translation of these developments. In addition to generally applicable principles for orthopedic study design, there are specific key criteria and considerations for cartilage repair studies. Based on a comprehensive literature review and standards of clinical care, the current report represents the discussions among scientific and clinical leaders, surgeons, and industry representatives and thus represents the ICRS consensus. It clearly identifies important scientific and clinical criteria for the design, reporting, and interpretation of studies in articular cartilage repair and provides general guidelines for the design and conduct of future study protocols in this field. Systematic application of these criteria and considerations will facilitate study designs that are scientifically rigorous, ethical, practical, and appropriate for the questions being addressed in any given cartilage repair research project.

Acknowledgments and Funding

The authors received no financial support for the research and/or authorship of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the authorship and/or publication of this article.

References


