THE ORTHOPAEDIC FORUM

Cartilage-Repair Innovation at a Standstill

Methodologic and Regulatory Pathways to Breaking Free

Stephen Lyman, PhD, Norimasa Nakamura, MD, PhD, Brian J. Cole, MD, MBA, Christoph Erggelet, MD, PhD, Andreas H. Gomoll, MD, and Jack Farr II, MD

Investigation performed at the Hospital for Special Surgery Healthcare Research Institute, New York, NY

Abstract: Articular cartilage defects strongly predispose patients to developing early joint degeneration and osteoarthritis, but for more than 15 years, no new cartilage-repair technologies that we know of have been approved by the U.S. Food and Drug Administration. Many studies examining novel approaches to cartilage repair, including cell, tissue, or matrix-based techniques, have shown great promise, but completing randomized controlled trials (RCTs) to establish safety and efficacy has been challenging, providing a major barrier to bringing these innovations into clinical use. In this article, we review reasons that surgical innovations are not well-suited for testing through RCTs. We also discuss how analytical methods for reducing bias, such as propensity scoring, make prospective observational studies a potentially viable alternative for testing the safety and efficacy of cartilage-repair and other novel therapies, offering the real possibility of therapeutic innovation.

Peer Review: This article was reviewed by the Editor-in-Chief and one Deputy Editor, and it underwent blinded review by 2 or more outside experts. It was also reviewed by an expert in methodology and statistics. The Deputy Editor reviewed each revision of the article, and it underwent a final review by the Editor-in-Chief prior to publication. Final corrections and clarifications occurred during one or more exchanges between the author(s) and copyeditors.

The repair of focal articular cartilage defects remains a clinical challenge, with few therapeutic options despite the well-established observation that damaged cartilage predisposes patients to premature osteoarthritis (OA). A variety of novel approaches, including cell, tissue, and matrix-based techniques, have shown great promise in laboratory or pilot clinical studies[1-6]. Since the approval of Carticel autologous chondrocytes for transplantation (Genzyme) in 1997[7], no new technologies for cartilage repair, to our knowledge, have been approved by the U.S. Food and Drug Administration (FDA). The major hurdle appears to be the completion of definitive randomized controlled trials (RCTs), with insufficient funding or enrollment undermining several recent attempts.

Although RCTs are currently the *conditio sine qua non* for FDA approval, there is no explicit statement that the required "substantial evidence" of safety and efficacy come from RCTs.

Disclosure: The authors indicated that no external funding was received for any aspect of this work. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article,* one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work and "yes" to indicate that the author had other relationships or activities that could be perceived to influence, or have the potential to influence, what was written in this work.

However, the FDA has offered little guidance with regard to alternative methodologies. Below, we review the hurdles for completing surgical RCTs and present alternative analytic approaches that improve the accuracy and rigor of prospective observational studies of innovative cartilage-repair therapies.

History of RCTs in Cartilage Repair

The 1962 Kefauver-Harris Drug Amendments established the FDA's authority in clinical study design and explicitly stated that drug or device approval would require substantial evidence of efficacy (Fig. 1). Although RCTs have been effective in demonstrating the efficacy and safety of new drugs, they pose particular challenges for evaluating novel surgical interventions, particularly when the condition is uncommon, as with cartilage injuries amenable to repair. Several attempts to complete Phase-III RCTs in cartilage repair have been delayed, halted, or moved outside the U.S. (Table I). DeNovo NT Natural Tissue (Zimmer) and BioCartilage (Arthrex) were recently approved as minimally treated grafts[8,9], but approvals of novel cell and matrix-based technologies have lagged. Low study enrollment suggests difficulties with recruitment and/or retention, which likely has contributed to the unsuccessful completion of trials, and there are other challenges, as discussed below.

Limitations of Surgical Cartilage-Repair RCTs

Cost

An RCT is the most expensive type of clinical study to successfully complete due to the infrastructure needed to comply with FDA standards, and surgical RCTs are particularly expensive. For example, the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)-funded Spine Patient Outcomes Research Trial (SPORT) assessing spine surgery versus conservative management enrolled 1,094 spine patients over 7 years at a cost of $21 million. Unfortunately, the study was ultimately flawed because of problems typical of surgical RCTs,

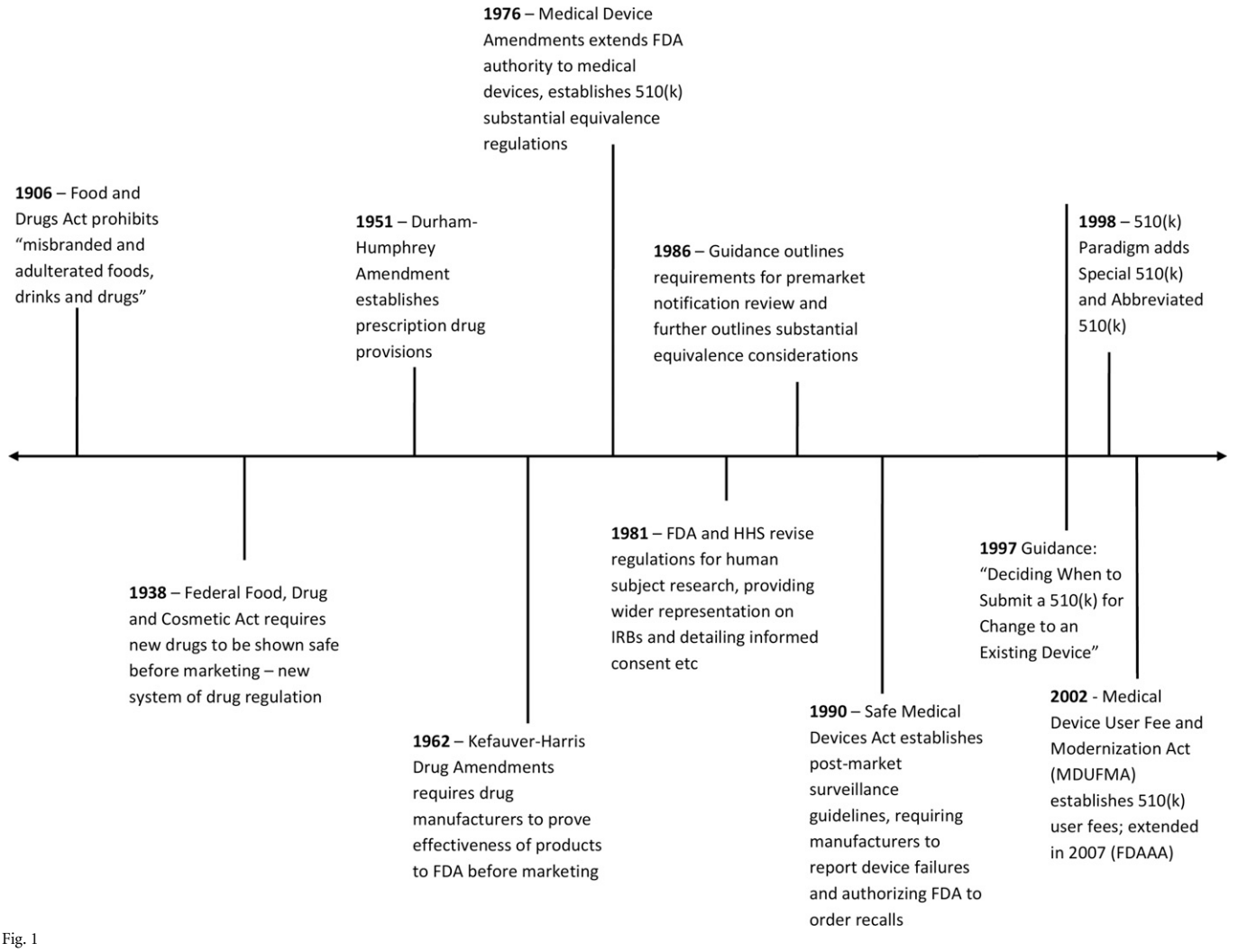

Fig. 1

A timeline detailing major events in the history of the U.S. Food and Drug Administration (FDA) and its regulatory guidelines for medical devices. HHS = U.S. Department of Health and Human Services, IRB = institutional review board, and FDAAA = FDA Amendments Act.

TABLE I Randomized Clinical Trials (RCTs) for Cartilage Repair*

Product/Technique	Company	Phase	Target Enrollment *(no.)*	Status	Date First Received	Date of Last Update
Cartilage Autograft Implantation System	DePuy Mitek	III	75	Completed	04/10/2009	02/26/2015
DeNovo NT engineered tissue†	Zimmer Orthobiologics	Safety/efficacy	200 (estimated)	Recruiting	04/01/2011	07/21/2015
Revaflex†	ISTO Technologies	III	225	Active, not recruiting	07/20/2011	03/23/2015
Osteofit	Kensey Nash	II	30	Terminated		
CarGel	BioSynTech Piramal Healthcare Canada	Observational	80	Enrolling (invitation only)	11/22/2010	09/21/2012
Matrix-induced chondrocyte implantation (MACI)	Sanofi	III	144	Completed	04/17/2008	08/17/2015
Carticel	Vericel (formerly Genzyme)	IV	126	Completed	09/08/2005	08/17/2015
ChondroCelect	TiGenix	III	118	Completed	12/21/2006	09/22/2011
BioCart II	ProChon Biotech	II	40	Unknown	08/04/2008	04/16/2012
NeoCart	Histogenics	III	245	Recruiting	02/08/2010	10/07/2015
CARTISTEM	Medipost	I, II	12	Active, not recruiting	11/20/2012	06/01/2015
NOVOCART 3D	Aesculap Biologics	III	233	Recruiting	09/17/2013	04/21/2015

*According to a search of ClinicalTrials.gov (http://clinicaltrials.gov/ct2/results?term=cartilage+repair+USA&Search=Search). Accessed 2016 Mar 10. †Same product, being tested under different names.

including low recruitment success and frequent treatment noncompliance and crossover (discussed below).

The high cost of surgical RCTs is due, in large part, to the procedure cost, because trial sponsors often pay for both experimental and control procedures. Most insurance companies preclude payment for procedures in an investigational study, and some institutional review boards object to patients bearing the cost of participation in investigational studies. In the ongoing NeoCart RCT, Histogenics is covering the costs of both the investigational treatment and the control microfracture surgery[10].

For companies developing cartilage-repair technologies, the cost of RCTs becomes particularly unappealing given the relatively small and crowded market. Isolated cartilage defects amenable to repair are not common, making them akin to an orphan condition despite the high risk of future OA. Introducing truly novel devices at prices reflecting the high cost of surgical RCTs is difficult when alternatives such as autografts or allografts are available at much lower prices, and less arduous approval pathways are available for minimally manipulated tissue and similar products. The many available options reflect the continued search for a clearly efficacious long-term solution[11-15]. Thus, market conditions and a challenging reimbursement environment support the need to explore alternatives to RCTs to establish safety and effectiveness of novel cartilage interventions.

Enrollment Challenges

The relatively small number of patients who qualify for an RCT in cartilage repair exacerbates recruitment challenges common to most surgical RCTs. The eligibility criteria are narrow, as only some injuries are amenable to repair. For example, in an ISTO Technologies study, patients must be 18 to 60 years old, with a BMI of <35 kg/m^2, and have 1 or 2 discrete cartilage lesions of ≤ 5 cm^2[2,16]. Strict inclusion and exclusion criteria are necessary in an RCT to limit heterogeneity and reduce confounding, but when few patients fit RCT inclusion criteria, the requirement for longer recruitment periods and/or more study sites increases cost. For example, the recently completed BST-CarGel trial required 26 sites and 10 years to successfully enroll and complete follow-up for just 80 patients[17].

Even when indications are broader, enrollment can be difficult, as recruiting a patient into an RCT comparing surgical and nonsurgical treatment is akin to asking, "Mind if we leave it up to chance whether or not you have surgery?" This can be vexing to patients, particularly those not familiar with the theoretical reasons for randomization. Even for patients who understand the potential societal benefits, an unwillingness to

leave the surgical decision to chance is understandable. The results of several recent orthopaedic RCTs provide evidence of this[18-22]. Patients may not accept randomization between any standard and experimental treatment, surgical or not, because they have ready access to information online, increasingly see medical care as a commodity[23,24], and are often looking for the "best" treatment. On the other hand, many patients still have limited health literacy and may not realize that they are not cartilage-restoration candidates. Engen et al. reported that, in a busy cartilage-restoration practice, nearly 90% of patients presenting for cartilage repair would not fit cartilage-repair RCT inclusion and/or exclusion criteria[25].

A further challenge is retention. In drug trials, efficacy can be demonstrated with relatively short follow-up, but cartilage-repair therapies often require 24 to 48 months of follow-up to show efficacy. Given that the U.S. population is more mobile than that of other countries, and employment changes and interstate migration are frequent[26], committing to long-term surgical follow-up is often not feasible, especially for eligible cartilage-repair patients, who tend to be younger adults.

The SPORT study successfully randomized only 27.5% of eligible patients[20]. An additional 35.5% consented to enrollment in a parallel observational cohort study, while the plurality (37.0%) declined participation entirely[20]. Similarly, the recently completed Meniscal Tear in Osteoarthritis Research (MeTeOR) trial comparing operative and nonoperative management of meniscal tears in patients with mild to moderate knee OA recruited only 26.4%[18,19]. The low rate of consent becomes particularly problematic for cartilage repair, wherein the number of eligible patients is already low. Of even greater concern is the limited generalizability of studies in which enrolled patients are a self-selected minority of the eligible population.

Noncompliance and Crossover

Compliance with treatment assignment is a common problem with drug trials, and methods have been devised to measure this compliance[27]. However, in surgical trials, compliance is especially problematic. In the SPORT study, 39.6% of the patients assigned to surgery did not have surgery within 2 years[20]. In the parallel observational cohort, 33.7% of the patients did not undergo surgery within 2 years, so patients left to their own choice underwent surgery more often than those assigned to surgery.

Meanwhile, crossover from nonoperative to operative management is so common that it can also undermine randomization. In the SPORT study, 44.0% of the nonoperative patients underwent surgery within 2 years, while 30.0% of nonoperative patients in the MeTeOR trial underwent surgery within 6 months[19,20]. Intent-to-treat and other statistical methods have been devised to account for noncompliance or crossover[27-30]. However, even these approaches cannot truly balance studies when rates at which patients change treatment groups are so high that they undermine the validity of randomization.

Equipoise

The ethical foundation of an RCT is *clinical* or *community equipoise,* which exists when uncertainty regarding the superiority of a particular treatment supports the need for an RCT. However, *individual equipoise* among participating surgeons is also necessary for the ethical conduct of an RCT. If a clinician believes that one treatment is superior to another, he/she cannot ethically deny that treatment on the basis of randomization.

The equipoise of an RCT is defined by the study investigators during trial design when establishing inclusion and exclusion criteria. For example, in a trial of knee debridement for mild to moderate knee OA, patients with large bucket-handle tears of the meniscus were excluded, as this is a clear indication for surgical intervention, regardless of OA status[31]. More generally speaking, with a surgical RCT, patients clearly indicated for operative or nonoperative management are excluded, with only patients of uncertain treatment preference being eligible for enrollment.

However, in a large multicenter surgical RCT, equipoise must also extend to each participating surgeon. While Surgeon A may believe that it is unethical to enroll patients whom he thinks are likely to benefit from surgery (Fig. 2), Surgeon B may believe it is unethical to enroll patients whom she thinks would be unlikely to benefit from surgery. Surgeon C could have treatment uncertainty that is narrower than the trial eligibility criteria, even if he is balanced in his assessment. All of these surgeons are likely to enroll only patients who are both eligible for the trial and match their individual equipoise. If a particular philosophy dominates among surgeon participants, or if a biased surgeon enrolls a high proportion of patients, the generalizability of the trial may be compromised (Fig. 3). A recent study by Katz et al. simulating an RCT on meniscal repair showed that surgeon preference could change the apparent efficacy of the treatment[32].

Pathway to Alternative Study Designs

Pressure from drug and device companies regarding the cost and difficulty of successfully completing RCTs has led the FDA

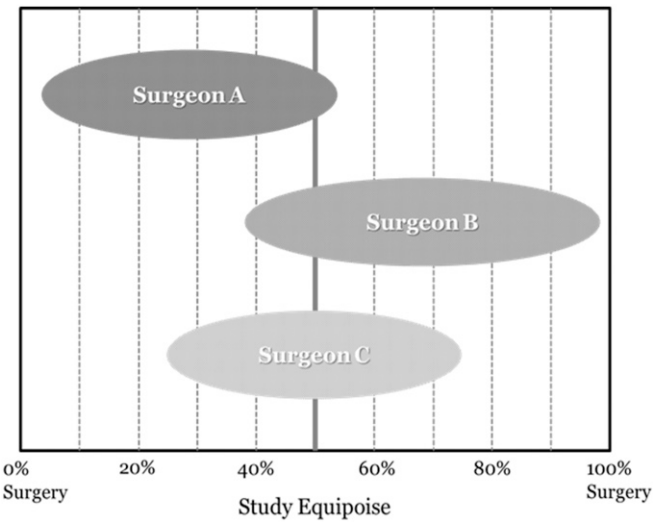

Fig. 2

Theoretical surgeon versus clinical equipoise in randomized controlled trials.

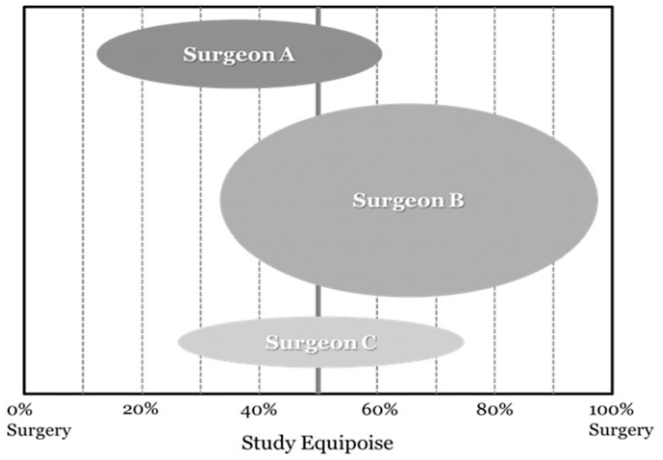

Fig. 3

Unbalanced surgeon equipoise and enrollment biases.

to allow the approval of drugs or devices with substantial equivalence to products already on the market, without clinical evidence, through a premarket notification, or 510(k), clearance process. However, novel cartilage repair technologies are not eligible for 510(k) clearance because no substantially equivalent therapies exist. Therefore, these technologies languish, because RCTs testing their efficacy have proven difficult or impossible to complete, precluding FDA approval (Table I). Despite robust evidence from experimental systems and pilot clinical studies, no novel cartilage-repair technology that we know of has been approved by the FDA in more than 15 years, emphasizing the necessity of developing alternative approval pathways for novel surgical interventions.

Alternatives to RCTs

The formation of the Patient-Centered Outcomes Research Institute (PCORI), authorized by the 2010 Patient Protection and Affordable Care Act, has given rise to the recognition of new approaches to clinical investigation. The 2013 *PCORI Methodology Report* established standards for prospective observational studies[33].

A well-designed, multicenter, prospective observational study represents a viable alternative to an RCT for determining surgical effectiveness. Prospective observational studies have relatively high recruitment success because patient treatment allocation is not randomly determined but rather, is decided through the physician-patient treatment decision-making process. Increased enrollment substantially improves the likelihood that study results will be broadly generalizable. Physician-patient decision-making obviates the problem of individual equipoise. Prospective observational studies are also less costly than RCTs, even after considering the application of costs to adhere to regulatory standards. Regulations to require the registration of prospective observational studies at ClinicalTrials.gov or similar sites could enhance transparency and increase the acceptance of these types of studies. In Japan, all cell-based clinical studies are conducted as single-arm studies, and the Pharmaceuticals and Medical Devices Agency of Japan recently announced regulations giving provisional approval to therapies that show probable benefit through such studies, thus facilitating commercialization[34].

Large, prospective, community-based registries of surgical interventions have been successfully established in cardiothoracic surgery and orthopaedics, building the infrastructure necessary to conduct high-quality, multicenter prospective observational studies. We strongly support the development of a large, international, multicenter registry for knee cartilage repair, data from which could help to establish safety, real-world effectiveness, and—with appropriate analyses—an estimate of efficacy of these therapies. With adequate governance from academic leaders, specialty societies, industry, and government representatives, thousands of patients undergoing surgical treatment for focal cartilage defects of the knee could be rapidly enrolled in such a registry. Recent reviews indicate a need for this kind of resource, with most clinical studies performed involving small cohorts[8,9,11-15]. A cartilage-repair registry would not only improve clinical research generally but would set the stage for an alternative approval pathway that is better suited to novel surgical therapies.

Prospective observational studies traditionally have been considered to have the second-highest level of evidence, behind RCTs, with selection bias considered the primary limitation. However, a historical review of observational studies on the use of hormone replacement therapy (HRT) exemplifies how new analytic methods can address this problem to provide more accurate information on safety and efficacy. In the early 2000s, results of RCTs by the Women's Health Initiative regarding postmenopausal use of HRT contradicted findings from the prospective cohort Nurses' Health Study[35]. Physicians prescribed HRT to millions of women for the cardioprotective effect reported in observational studies, only to discover that subsequent RCTs showed that HRT put women at risk of cardiovascular disease. However, a reanalysis of the Women's Health Initiative observational HRT cohort using propensity score matching[36] produced results nearly identical to those of the RCTs: HRT increased the risk of cardiovascular events. This series of studies exemplifies how thoughtful adjustments to an analytic approach can reduce biases in prospective observational studies.

Selection bias is highly relevant when considering surgical interventions for cartilage repair. Depending on the size of the lesion, treatment options for cartilage defects may include nonoperative therapy, debridement, microfracture or other drilling techniques, or osteochondral autografting or allografting, in addition to use of the emerging cell or matrix-augmented techniques. While cartilage repair is intended to forestall degeneration of the affected joint by recreating a more normal anatomy and kinematics, some patients may opt for nonoperative treatment or debridement to avoid a lengthy recovery and return to full activity more quickly. While these types of patient or surgeon choices could, in a traditional analysis of observational study data, yield spurious results, newer analytic techniques eliminate or minimize bias.

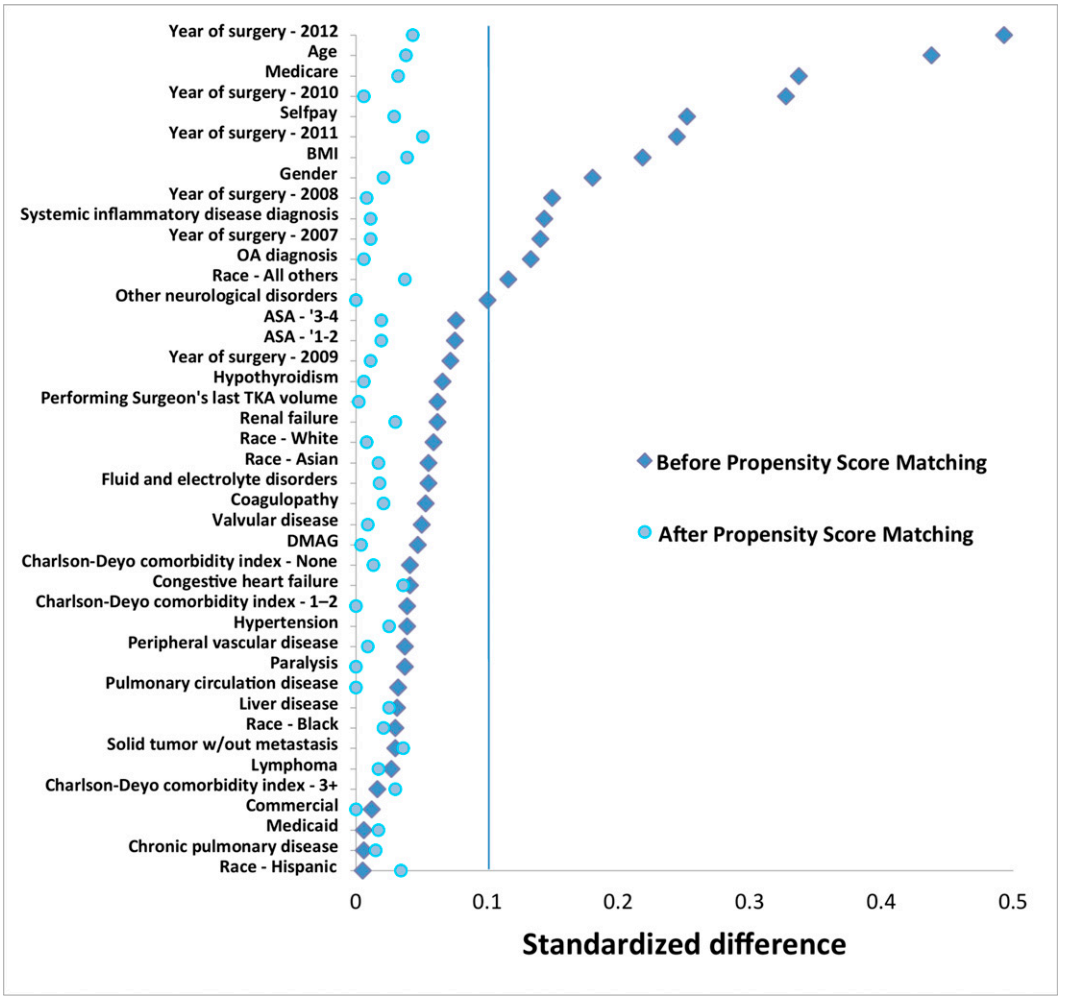

Fig. 4

Propensity score matching minimizes differences between measured covariates in a sample cohort. Standardized differences were calculated for the 42 variables (along the y axis) for which patients in a sample cohort undergoing 2 types of knee procedures were matched. The standardized difference (also known as the Cohen effect size index, along the x axis) compares 2 sample means in units of the pooled standard deviation so that a standardized difference of ≥0.1 denotes meaningful imbalance in the variable. In this example, there is an imbalance for 14 variables among patients in the 2 treatment groups before propensity score matching. However, after propensity score matching, the likelihood that these variables affected each patient's assignment to a particular treatment was reduced.

We believe that these new methods improve the ability of prospective observational studies to serve as an alternative to RCTs for evaluating new cartilage-repair techniques or other surgical procedures. The *PCORI Methodology Report* specifically recommended 2 analytic techniques to reduce or eliminate the types of bias inherent in observational study designs: the instrumental variable and propensity scoring[33].

Instrumental Variable

An instrumental variable can be a useful tool when there is a strong preference toward one therapy over another that may not be based on evidence but rather, on the physician's or patient's preference. A "good" instrumental variable adjusts for bias by being predictive of patient treatment allocation but is unrelated to patient outcome. For example, a recent trial evaluating anesthesia choice during hip fracture surgery used the distance from a patient's residence to a hospital that uses regional anesthesia in the treatment of hip fractures as an instrumental variable because patient access to a center that uses regional anesthesia would influence the likelihood of a patient receiving regional anesthesia but would be unrelated to anesthesia outcomes[37]. This may serve as an appropriate instrumental variable in the context of surgical studies as well.

Propensity Scores

Propensity scores are generated from regression models that include all potentially important patient characteristics that could influence a clinician's treatment decision. Thus, the model estimates treatment allocation, rather than the outcome. Model coefficients are used to assign each patient a propensity score

ranging from 0 to 1 to represent the patient's likelihood of receiving treatment A versus treatment B. A score of 0 indicates no chance of receiving treatment A, while a score of 1 represents absolute certainty of receiving treatment A. A score of 0.5 would represent a virtual coin flip in treatment choice.

Once scores are calculated for each patient, they can be used in a variety of ways. For example, a propensity score "match" could be performed, which replicates randomization. In a study of cartilage repair, a patient with a propensity score of 0.234 for microfracture who underwent microfracture surgery would be matched with a patient from the alternative cartilage-treatment cohort having a propensity score nearest to 0.234. A variety of methods exist to determine which matches are best, but when a match is appropriately performed, clinical characteristics of the 2 treatment groups will be balanced, minimizing or eliminating the previously existing selection bias in the cohort. This can be visualized by plotting the standardized differences for each variable before and after matching (Fig. 4). A standardized difference is employed to assess balance for each variable that was used to estimate the propensity score. It was proposed in 1962 as the Cohen effect size index for the comparison of 2 sample means in units of the pooled standard deviation[38,39]. A standardized difference of ≥0.1 denotes meaningful imbalance for the variable.

Whereas randomization eliminates bias even from unknown confounders, in a strict sense, propensity matching can only account for known factors. However, because the matching process balances all measured covariates, the effects of unmeasured confounders may also be minimized. Despite this, it behooves researchers in cartilage repair to develop a standardized list of parameters that best represents the most important baseline characteristics for creating a propensity score. Propensity score matching has not, to our knowledge, been used previously in arthroscopic surgery research, but it has been used in other areas of orthopaedics, including arthroplasty and trauma[40,41].

A potential drawback of propensity score matching is that a very large number of patients may be needed, especially in the untreated group[42]. Moreover, matching frequently omits a substantial proportion of the population when comparison groups are being constructed, which is problematic if unmatched patients had a rare but informative outcome (e.g., revision surgery). Therefore, inverse probability of treatment weighting (IPTW) is proposed as an alternative to matching to adjust for confounding[43-45]. With IPTW, the propensity score is used as a weight in a weighted regression. The inverse of the propensity score is used to weight each observation in the treated group, and 1 minus the inverse of the propensity score is used to weight each observation in the untreated group. Weights restore balance between the clinical characteristics of the 2 treatment groups and allow for the use of the entire sample, rather than the subset of matched patients.

Overcoming the Challenges

Practical barriers limit successful completion of surgical RCTs even when the evaluated surgical techniques are established clinical practice. The continued failure to complete RCTs involving cartilage-regeneration therapies shows that these challenges are even greater when testing novel surgical interventions. While RCTs remain a gold standard in many cases, analytic methods that reduce bias can provide additional rigor to make prospective observational studies viable alternatives for evaluating the safety and efficacy of cartilage-repair and other novel therapies, offering the real possibility of therapeutic innovation. ■

Stephen Lyman, PhD[1]
Norimasa Nakamura, MD, PhD[2]
Brian J. Cole, MD, MBA[3]
Christoph Erggelet, MD, PhD[4]
Andreas H. Gomoll, MD[5]
Jack Farr II, MD[6]

[1]Hospital for Special Surgery Healthcare Research Institute, New York, NY

[2]Department of Orthopedics, Osaka Health Science University, Osaka, Japan

[3]Department of Orthopedics, Rush University Medical Center, Chicago, Illinois

[4]Center for Biologic Joint Surgery, Department of Orthopaedic Surgery and Traumatology, University of Freiburg Medical Center, Freiburg, Germany

[5]Department of Orthopaedic Surgery, Brigham & Women's Hospital, Harvard Medical School, Boston, Massachusetts

[6]Indiana University School of Medicine and OrthoIndy Hospital, Indianapolis, Indiana

E-mail address for S. Lyman: lymans@hss.edu

References

1. Hui JH, Goyal D, Nakamura N, Ochi M; Asian Cartilage Society. Cartilage repair: 2013 Asian update. Arthroscopy. 2013 Dec;29(12):1992-2000.

2. Johnstone B, Alini M, Cucchiarini M, Dodge GR, Eglin D, Guilak F, Madry H, Mata A, Mauck RL, Semino CE, Stoddart MJ. Tissue engineering for articular cartilage repair—the state of the art. Eur Cell Mater. 2013;25:248-67. Epub 2013 May 2.

3. Marcacci M, Filardo G, Kon E. Treatment of cartilage lesions: what works and why? Injury. 2013 Jan;44(Suppl 1):S11-5.

4. Moran CJ, Pascual-Garrido C, Chubinskaya S, Potter HG, Warren RF, Cole BJ, Rodeo SA. Restoration of articular cartilage. J Bone Joint Surg Am. 2014 Feb 19;96(4):336-44.

5. Roelofs AJ, Rocke JP, De Bari C. Cell-based approaches to joint surface repair: a research perspective. Osteoarthritis Cartilage. 2013 Jul;21(7):892-900. Epub 2013 Apr 15.

6. Ye K, Felimban R, Moulton SE, Wallace GG, Di Bella C, Traianedes K, Choong PF, Myers DE. Bioengineering of articular cartilage: past, present and future. Regen Med. 2013 May;8(3):333-49.

7. Genzyme. Carticel. 2007 Jun. http://www.fda.gov/downloads/BiologicsBloodVaccines/CellularGeneTherapyProducts/ApprovedProducts/UCM109339.pdf. Accessed 2016 Feb 1.

8. Tompkins M, Adkisson DH, Bonner KF. DeNovo NT allograft. Oper Tech Sports Med. 2013;21:82-9.

9. Abrams GD, Mall NA, Fortier LA, Roller BL, Cole BJ. BioCartilage: background and operative technique. Oper Tech Sports Med. 2013;21:116-24.

10. Crawford DC, DeBerardino TM, Williams RJ 3rd. NeoCart, an autologous cartilage tissue implant, compared with microfracture for treatment of distal femoral cartilage lesions: an FDA phase-II prospective, randomized clinical trial after two years. J Bone Joint Surg Am. 2012 Jun 6;94(11):979-89.

11. Kon E, Roffi A, Filardo G, Tesei G, Marcacci M. Scaffold-based cartilage treatments: with or without cells? A systematic review of preclinical and clinical evidence. Arthroscopy. 2015 Apr;31(4):767-75. Epub 2015 Jan 27.

12. Oussedik S, Tsitskaris K, Parker D. Treatment of articular cartilage lesions of the knee by microfracture or autologous chondrocyte implantation: a systematic review. Arthroscopy. 2015 Apr;31(4):732-44. Epub 2015 Feb 3.

13. Wylie JD, Hartley MK, Kapron AL, Aoki SK, Maak TG. What is the effect of matrices on cartilage repair? A systematic review. Clin Orthop Relat Res. 2015 May;473(5):1673-82.

14. Wylie JD, Hartley MK, Kapron AL, Aoki SK, Maak TG. Failures and reoperations after matrix-assisted cartilage repair of the knee: a systematic review. Arthroscopy. 2016 Feb;32(2):386-92. Epub 2015 Sep 28.

15. Saris D, Price A, Widuchowski W, Bertrand-Marchand M, Caron J, Drogset JO, Emans P, Podskubka A, Tsuchida A, Kili S, Levine D, Brittberg M; SUMMIT study group. Matrix-applied characterized autologous cultured chondrocytes versus microfracture: two-year follow-up of a prospective randomized trial. Am J Sports Med. 2014 Jun;42(6):1384-94. Epub 2014 Apr 8.

16. Clinicaltrials.gov. Neocartilage implant to treat cartilage lesions of the knee. https://clinicaltrials.gov/ct2/show/NCT01400607?term=cartilage+repair+USA&rank=5. Accessed 2016 Feb 1.

17. Shive M, Stanish WD, McCormack RG, Mohtadi N, Pelet S, Desnoyers J, Forriol F, Vehik K, Methot S, Restrepo A. BST-CarGel treatment results in sustained cartilage repair superiority compared to microfracture at 5 years. Read at the 10th Biennial ISAKOS Congress. 2015 Jun 7-11; Lyon, France. Paper no. 65.

18. Katz JN, Brophy RH, Chaisson CE, de Chaves L, Cole BJ, Dahm DL, Donnell-Fink LA, Guermazi A, Haas AK, Jones MH, Levy BA, Mandl LA, Martin SD, Marx RG, Miniaci A, Matava MJ, Palmisano J, Reinke EK, Richardson BE, Rome BN, Safran-Norton CE, Skoniecki DJ, Solomon DH, Smith MV, Spindler KP, Stuart MJ, Wright J, Wright RW, Losina E. Surgery versus physical therapy for a meniscal tear and osteoarthritis. N Engl J Med. 2013 May 2;368(18):1675-84. Epub 2013 Mar 18.

19. Katz JN, Chaisson CE, Cole B, Guermazi A, Hunter DJ, Jones M, Levy BA, Mandl LA, Martin S, Marx RG, Safran-Norton C, Roemer FW, Skoniecki D, Solomon DH, Spindler KP, Wright J, Wright RW, Losina E. The MeTeOR trial (Meniscal Tear in Osteoarthritis Research): rationale and design features. Contemp Clin Trials. 2012 Nov;33(6):1189-96. Epub 2012 Sep 5.

20. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, Abdu WA, Hilibrand AS, Boden SD, Deyo RA. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. JAMA. 2006 Nov 22;296(20):2441-50.

21. Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, Herkowitz H, Fischgrund J, Cammisa FP, Albert T, Deyo RA. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. JAMA. 2006 Nov 22;296(20):2451-9.

22. Weinstein JN, Lurie JD, Tosteson TD, Zhao W, Blood EA, Tosteson AN, Birkmeyer N, Herkowitz H, Longley M, Lenke L, Emery S, Hu SS. Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. Four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. J Bone Joint Surg Am. 2009 Jun;91(6):1295-304.

23. Lee CJ, Gray SW, Lewis N. Internet use leads cancer patients to be active health care consumers. Patient Educ Couns. 2010 Dec;81(Suppl):S63-9.

24. Michalak J, Schreiner NJ, Tennis W, Szekely L, Hale M, Guastello S. 'The patient will see you now'. Am J Nurs. 2010 Jan;110(1):61-3.

25. Engen CN, Engebretsen L, Årøen A. Knee cartilage defect patients enrolled in randomized controlled trials are not representative of patients in orthopedic practice. Cartilage. 2010 Oct;1(4):312-9.

26. Ihrke, D. Reason for moving: 2012 to 2013: population characteristics. 2014 Jun. https://www.census.gov/prod/2014pubs/p20-574.pdf. Accessed 2016 Feb 1.

27. Farmer KC. Methods for measuring and monitoring medication regimen adherence in clinical trials and clinical practice. Clin Ther. 1999 Jun;21(6):1074-90; discussion 1073.

28. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. Stat Med. 1997 May 15;16(9):1017-29.

29. Nagelkerke N, Fidler V, Bernsen R, Borgdorff M. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. Stat Med. 2000 Jul 30;19(14):1849-64.

30. Weinstein GS, Levin B. Effect of crossover on the statistical power of randomized studies. Ann Thorac Surg. 1989 Oct;48(4):490-5.

31. Kirkley A, Birmingham TB, Litchfield RB, Giffin JR, Willits KR, Wong CJ, Feagan BG, Donner A, Griffin SH, D'Ascanio LM, Pope JE, Fowler PJ. A randomized trial of arthroscopic surgery for osteoarthritis of the knee. N Engl J Med. 2008 Sep 11;359(11):1097-107.

32. Katz JN, Wright J, Levy BA, Baron JA, Losina E. Departures from community equipoise may lead to incorrect inference in randomized trials. J Clin Epidemiol. 2011 Mar;64(3):280-5. Epub 2010 Jul 17.

33. Hickam D, Totten A, Berg A, Rader K, Goodman S, Newhouse R, editors. The PCORI methodology report. 2013 Nov. http://www.pcori.org/assets/2013/11/PCORI-Methodology-Report.pdf. Accessed 2015 Jul 6.

34. Ogura T, Tsuchiya A, Mizuno S. Ongoing studies of cell-based therapies for articular cartilage defects in Japan. Orthopaedic Research and Reviews. 2015;7:1-10.

35. Prentice RL, Pettinger M, Anderson GL. Statistical issues arising in the Women's Health Initiative. Biometrics. 2005 Dec;61(4):899-911; discussion 911–41.

36. Tannen RL, Weiner MG, Xie D, Barnhart K. A simulation using data from a primary care practice database closely replicated the Women's Health Initiative trial. J Clin Epidemiol. 2007 Jul;60(7):686-95. Epub 2006 Dec 27.

37. Neuman MD, Rosenbaum PR, Ludwig JM, Zubizarreta JR, Silber JH. Anesthesia technique, mortality, and length of stay after hip fracture surgery. JAMA. 2014 Jun 25;311(24):2508-17.

38. Flury BK, Riedwyl H. Standard distance in univariate and multivariate analysis. Am Stat. 1986;40:249-51.

39. Normand ST, Landrum MB, Guadagnoli E, Ayanian JZ, Ryan TJ, Cleary PD, McNeil BJ. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: a matched analysis using propensity scores. J Clin Epidemiol. 2001 Apr;54(4):387-98.

40. Duchman KR, Gao Y, Pugely AJ, Martin CT, Callaghan JJ. Differences in short-term complications between unicompartmental and total knee arthroplasty: a propensity score matched analysis. J Bone Joint Surg Am. 2014 Aug 20;96(16):1387-94.

41. Jenkinson RJ, Kiss A, Johnson S, Stephen DJ, Kreder HJ. Delayed wound closure increases deep-infection rate associated with lower-grade open fractures: a propensity-matched cohort study. J Bone Joint Surg Am. 2014 Mar 5;96(5):380-6.

42. Guo S, Fraser MW. Propensity score analysis: statistical methods and applications. 2nd ed. Thousand Oaks: Sage Publications; 2014 Jul 11.

43. Cassel CM, Sarndal CE, Wretman JH. Some uses of statistical models in connection with the nonresponse problem. In: Madow WG, Olkin I, editors. Incomplete data in sample Surveys, vol. 3: proceedings of the symposium. New York: Academic Press; 1983.

44. Hirano K, Imbens GW. Estimation of causal effects using propensity score weighting: an application to data on right heart catheterization. Health Serv Outcomes Res Methodol. 2001;2:259-78.

45. Rosenbaum PR. Model-based direct adjustment. J Am Stat Assoc. 1987;82:387-94.