

Donor-Site Burden and Multidomain Outcomes in Comparative Graft Studies

Simon Crawford, Vincent Barker, Dominic Warren

Abstract

A graft comparison can improve one outcome while worsening another, especially when obtaining the graft introduces an additional anatomical site into the patient's recovery. This conceptual synthesis develops a multidomain framework for evaluating reconstructive strategies. Superior capsular graft biology and adult revision ACL management motivate separate applications. The framework distinguishes index-site function, donor-site burden, structural events, additional treatment, and patient priorities. It explains why a favorable result on one domain does not determine net benefit and why composite outcomes require explicit weighting and ascertainment. Established shoulder and knee instruments are used as examples of domain-specific measurement, while causal and competing-risk methods clarify the interpretation of longitudinal comparisons. No new patient scores, treatment effects, or utility weights are estimated. The article contributes a prospective outcome design that preserves tradeoffs and makes the basis of an overall judgment visible.

Keywords: donor-site burden; graft outcomes; patient priorities; multidomain assessment; comparative effectiveness.

1. A second anatomical site changes the comparison

Evaluating a reconstruction only at the index joint can omit part of the treatment experience. When graft acquisition involves another anatomical site, symptoms or limitations there belong to the strategy's consequences. An alternative strategy can avoid that acquisition while introducing different biological or logistical considerations. The scientific comparison should therefore identify the whole intervention and its outcome domains. Otherwise a study can appear to establish superiority by measuring the domain most favorable to one strategy and leaving the competing burden outside the definition of success.

Cao and colleagues' superior capsular reconstruction review motivates considering biological differences among graft options (Cao et al., 2025). Their APKASS revision ACL statement motivates a different clinical decision context (Cao et al., 2026). The present framework connects those contexts through the structure of a comparison, not through an assumed common effect. Shoulder and knee procedures are not pooled, and no graft is recommended for a particular patient. The question is how a study should preserve the benefits and burdens that would be needed to evaluate a strategy fairly within either setting.

The article is a conceptual synthesis. It does not estimate donor-site event rates or infer patient preferences from published averages. Its starting proposition is that an overall judgment requires a visible set of outcomes and a stated method for relating them. A biological advantage can be a meaningful reason to investigate a graft while remaining insufficient to determine the balance of patient consequences. Keeping the domains separate allows a study to show where that balance is known, where it is uncertain, and which additional observation would be useful.

2. Define the domains before collecting scores

Index-site symptoms, function, and structural status are not redundant outcomes. Pain can change without an equivalent change in participation, and an imaging finding can differ from the patient's experience. A domain map should identify which aspect of recovery each measurement represents. The Shoulder Pain and Disability Index supplies an established example of a shoulder-focused instrument (Roach et al., 1991). Its relevance is the explicit content of the construct being measured. A score should not be described as overall health or complete graft success merely because it is familiar and easy to administer.

Knee-focused instruments provide different content. The KOOS measures specified domains related to knee injury and osteoarthritis (Roos et al., 1998). The IKDC subjective knee form provides another defined knee outcome framework (Irrgang et al., 2001). Investigators should choose an instrument because its content matches the study question and population, not because one scale produced the most favorable result in earlier reports. Using both may be informative, but it introduces additional outcomes whose role should be specified. Their numerical units should not be treated as interchangeable with a shoulder score or with a donor-site measurement.

Donor-site assessment needs its own anatomical and functional support. A general index-joint questionnaire may not ask about the additional site or the activity affected by it. The study should identify whether the burden is transient or persistent, how it was elicited, and whether the participant could attribute it confidently. An open symptom report can complement a structured measure without automatically supplying a causal attribution. Recording the timing and character of the complaint helps distinguish an event after graft acquisition from an event demonstrated to have been caused by it.

3. The complete strategy is the intervention

A comparison between graft strategies may include acquisition, processing, fixation, and subsequent care. If the research question concerns that complete pathway, donor-site consequences should remain part of the outcome rather than be adjusted away. If the question concerns a narrower material effect under otherwise standardized conditions, another design may be required. Causal diagrams help state which variables precede selection and which lie on a treatment pathway (Greenland et al., 1999). The analytical choice should follow this specification instead of being determined by which adjustment makes the main coefficient appear favorable.

Selection can complicate donor-site comparisons. Patients receiving one strategy may differ in prior procedures, available tissue, activity expectations, or baseline symptoms at the potential acquisition site. A later group difference can therefore reflect both treatment and the conditions that influenced its selection. The baseline record should include the measurements needed for the proposed contrast. A blank donor-site field before treatment makes it difficult to distinguish a newly developed burden from a preexisting limitation. Detailed outcome collection cannot fully repair an absent baseline relevant to the question.

The MARS cohort illustrates a clinical research setting in which graft choice and subsequent revision-ACL outcomes can be examined (MARS Group, 2014). Its role here is to emphasize that graft selection is embedded in a broader clinical process. This article does not extract a new donor-site effect from that cohort or assume that its findings transport to superior capsular reconstruction. A prospective study within either setting would need its own treatment definitions, baseline measurements, outcome schedule, and approach to confounding. The whole comparison should remain visible when a material-specific explanation is discussed.

4. Tradeoffs cannot be resolved by significance alone

Suppose one strategy has a favorable estimate for index-site function and an unfavorable estimate for another burden. Statistical significance on one side does not determine the overall balance. The two estimates may differ in precision, scale, duration, and importance to patients. A report should show the absolute outcome distributions and their uncertainty rather than compare p values as if they were measures of benefit. A domain with a wide interval is not necessarily unimportant, and a small precisely estimated difference is not automatically decisive.

An outcome profile can make the tradeoff explicit. It would retain the primary domain, prespecified secondary domains, timing, and relevant adverse events. The profile should not become a menu from which investigators choose a favorable conclusion after inspection. A hierarchy of questions and a clear distinction between confirmatory and exploratory findings can preserve interpretability. Where patient priorities differ, reporting the profile may be more honest than forcing one universal ranking. The present article proposes that reporting structure without inventing weights or claiming that every patient values the domains in the same way.

Evidence certainty also matters. GRADE separates judgments about evidence from the strength of a recommendation (Guyatt et al., 2008). A comparison with compelling biological reasoning but uncertain clinical estimates may justify further research rather than a strong overall ranking. Conversely, a consistent patient-relevant effect can remain difficult to explain mechanistically. The outcome profile should identify both the direction of the observed results and the certainty supporting them. Agreement among experts can inform which tradeoffs need study, but it cannot supply unobserved patient preferences or convert an uncertain effect into a precise one.

5. Composite outcomes and patient priorities

A composite endpoint can be convenient when several events matter, but its components need a coherent interpretation. Combining persistent donor-site symptoms, an imaging event, and repeat surgery into one binary outcome gives each component an implicit role. The most frequent or most intensively observed component can dominate the result. Investigators should report component frequencies and explain the rationale for the combination. A favorable composite does not establish an improvement in every component, and it can conceal a tradeoff if some events become less common while others increase.

Patient preference measurement can make those judgments more explicit, but it requires its own design. Preferences may vary with activity goals, prior experiences, and the way information is presented. A hypothetical choice made before treatment is different from a retrospective satisfaction response after recovery. Both can be useful when their timing and meaning are preserved. The study should avoid assigning preferences from demographic categories alone or assuming that a clinician's ranking is identical to the patient's. This article does not perform a preference survey or estimate a utility function.

Time is part of the balance. A transient burden and a persistent limitation should not be treated as identical merely because both occurred at least once. Repeated measurement can distinguish onset, duration, and resolution. The schedule should be fine enough to capture the process relevant to the outcome without imposing unsustainable burden on participants. A final-visit score can miss a substantial early experience, while a cumulative event indicator can exaggerate a resolved minor problem. Reporting both trajectory and final state can help when they answer prespecified complementary questions.

6. Observation and competing events

Additional procedures can alter both the index-site outcome and the donor-site trajectory. A participant may undergo treatment elsewhere, leaving the original study with incomplete information. Event verification and permitted record linkage can help, but their availability can differ across groups. A study should distinguish no observed event from confirmed absence of an event. The difference is especially important when an apparently lower adverse-event rate can arise from less complete surveillance. A common contact and ascertainment plan strengthens interpretation of the entire outcome profile.

Competing-risk methods address defined event questions when other events change the opportunity for the event of interest (Fine and Gray, 1999). For example, a major subsequent intervention can alter which later graft-related outcomes are meaningful. The choice of framework depends on the estimand and the event process. It should not be introduced simply because multiple outcomes are available. The data should retain event order and reasons for ending follow-up so that the analysis can distinguish a treatment transition from ordinary loss of observation. No event model is fitted in this article.

Multiplicity should be considered when numerous domains, visits, and subgroups generate inferential comparisons. False discovery rate procedures concern a specified family of tests under stated assumptions (Benjamini and Hochberg, 1995). They do not decide which outcome is most meaningful or remove bias from selective reporting. A stronger design identifies primary questions, preserves the complete tested set, and treats exploratory patterns as candidates for later confirmation. Correcting only the favorable reported comparisons would miss the selection process. Transparent outcome planning is therefore part of the evidence, not a cosmetic addition to the final table.

7. Designing the next comparison

The distribution of burden can matter as much as its average. A modest mean difference can conceal a small subgroup with persistent difficulty, while a frequent short-lived complaint can dominate an event count without describing long-term function. Reporting relevant quantiles, persistence, and the proportion with unresolved symptoms can complement a mean when the sample supports those descriptions. The analysis should retain the uncertainty in small subgroups and avoid creating a threshold solely to produce a striking comparison. A multidomain study can therefore improve interpretation without inventing one universal utility score. Its value is to show which experiences are common, which persist, and which remain too sparsely observed for a precise judgment.

A prospective record could begin with a baseline map of index-site and potential donor-site status, participant goals, and the reasons for strategy selection. It would then follow a compact common set of domains at declared intervals and record additional treatment consistently. The analysis plan would distinguish the complete strategy effect from narrower mechanistic questions. It would also state which tradeoffs can be described directly and which require preference information not collected by the study. This design makes the missing observations visible before a single overall conclusion is attempted.

The biological questions motivating the superior capsular review and the management questions motivating the ACL revision statement can both contribute to that planning. Their joint value is to broaden the comparison beyond one favorable endpoint while preserving each procedure's context. Neither source establishes a universal weighting of donor-site burden against other outcomes. A future study should test the relevant domains within its own eligible population and make any synthesis rule explicit. Cross-procedure learning can then concern the measurement strategy rather than an unsupported transfer of treatment effects.

Comparative graft research should make its benefits and burdens jointly visible. Patient-reported function, structural observations, additional procedures, and donor-site experiences can provide complementary evidence without yielding one automatic ranking. A complete strategy definition, appropriate baseline data, and a prespecified outcome profile make that evidence easier to interpret. The framework developed here offers those design elements and leaves the overall balance to studies with the necessary outcome and preference observations. It reports no new superiority claim and no new estimate of patient burden.

References

- Cao, Mingde, et al. "Superior capsule reconstruction graft selection: The influence of biological properties of grafts on healing and re-tearing." *Bioengineering* 12.9 (2025): 942.
- Cao, Mingde, et al. "APKASS 2024 consensus statement on anterior cruciate ligament reconstruction, Part II: Management of anterior cruciate ligament revision in adults." *Asia-Pacific Journal of Sports Medicine, Arthroscopy, Rehabilitation and Technology* 45 (2026): 74-81.
- The MARS Group. "Effect of Graft Choice on the Outcome of Revision Anterior Cruciate Ligament Reconstruction in the Multicenter ACL Revision Study (MARS) Cohort." *The American Journal of Sports Medicine*, vol. 42, no. 10, 2014, pp. 2301-2310. <https://doi.org/10.1177/0363546514549005>.
- Roach, Kathryn E., et al. "Development of a Shoulder Pain and Disability Index." *Arthritis & Rheumatism*, vol. 4, no. 4, 1991, pp. 143-149. <https://doi.org/10.1002/art.1790040403>.
- Roos, Ewa M., et al. "Knee Injury and Osteoarthritis Outcome Score (KOOS): Development of a Self-Administered Outcome Measure." *Journal of Orthopaedic & Sports Physical Therapy*, vol. 28, no. 2, 1998, pp. 88-96. <https://doi.org/10.2519/jospt.1998.28.2.88>.
- Irrgang, James J., et al. "Development and Validation of the International Knee Documentation Committee Subjective Knee Form." *The American Journal of Sports Medicine*, vol. 29, no. 5, 2001, pp. 600-613. <https://doi.org/10.1177/03635465010290051301>.
- Fine, Jason P., and Robert J. Gray. "A Proportional Hazards Model for the Subdistribution of a Competing Risk." *Journal of the American Statistical Association*, vol. 94, no. 446, 1999, pp. 496-509. <https://doi.org/10.1080/01621459.1999.10474144>.
- Benjamini, Yoav and Yosef Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 57, no. 1, 1995, pp. 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Greenland, Sander, Judea Pearl, and James M. Robins. "Causal Diagrams for Epidemiologic Research." *Epidemiology*, vol. 10, no. 1, 1999, pp. 37-48. <https://doi.org/10.1097/00001648-199901000-00008>.
- Guyatt, Gordon H, et al. "GRADE: an emerging consensus on rating quality of evidence and strength of recommendations." *BMJ*, vol. 336, no. 7650, 2008, pp. 924-926. <https://doi.org/10.1136/bmj.39489.470347.ad>.