

Finding Joy in the Practice of Implementation Science: What Can We Learn from a Negative Study?

Lisa V. Rubenstein, MD, MSPH

University of California Los Angeles David Geffen School of Medicine and Fielding School of Public Health, RAND Corporation, Santa Monica, CA, USA.

J Gen Intern Med 34(1):9–11
DOI: 10.1007/s11606-018-4715-0
© Society of General Internal Medicine 2018

The practice of implementation science is the ongoing hard work of translating scientific learning into actualized and measured improvements in healthcare delivery. Implementers integrate delivery system data and know-how with findings from scientific literature and carry out and study their interventions within the system's management structures. Studies testing the effects of these interventions fall within the broad category of quality improvement but address the specific case of delivery system implementation of scientifically based interventions. The implemented interventions are not usually based on a single study, but rather reflect synthesis, such as systematic review-based recommendations or guidelines. The study by McWilliams and colleagues, entitled "Aiming to Improve Readmissions Through InteGrated Hospital Transitions" (AIRTIGHT) and published in *JGIM*,¹ is an example of a high-quality, negative quality improvement intervention (QII)² study aimed at implementing research-based recommendations. The study has much to teach both about reducing readmissions and about the practice of implementation science.

In terms of reducing readmissions, the AIRTIGHT study calls into question the adequacy or reliability across delivery settings of current systematic review-based recommendations for reducing readmissions. The study identified hospitalized patients at high risk of readmission (high-risk patients) using a computerized risk index, randomized them, and initiated referral to a patient navigator for those allocated to the intervention group. The navigator alerted a transition team at discharge. The team carried out weekly phone or in person follow-up for 30 days post discharge. Follow-up focused on MD clinical assessment, medication reconciliation, care management, and in some cases home visits. The study followed the basic recommendation-based principles for engaging a dedicated interdisciplinary team across inpatient and outpatient settings. Nevertheless, the study did not show achievement of its primary outcome of reduced readmissions across the patients it intended to treat. Even a matched subgroup

analysis of intervention group patients comparing those who received and did not receive transition showed no effect on readmission. We can confidently conclude that the study intervention, as carried out in this study, did not achieve its intended outcomes.

The AIRTIGHT study's use of rigorous comparison is noteworthy. While the literature on reducing acute care use among complex or high-risk patient interventions abounds with quality improvement studies showing a pre-post reduction in acute care use, often around 20%, most of these studies do not include a randomly assigned control group. A pre-post difference in acute care use, such as the approximately 15% drop among both intervention and control patients in AIRTIGHT, is especially likely to be due to regression to the mean when subjects are high-risk patients. These patients are acute care use outliers at the outset; we can expect their acute care use to normalize over time. The randomized comparison used to test AIRTIGHT may thus have helped the Carolinas HealthCare System, which funded the study, avoid false optimism about the intervention and eschew substantial downstream costs for maintaining or disseminating it.

Understanding how a QII was implemented is critical for making use of evaluation findings and begins with learning about the intervention's intended aims, rationale, and design.^{3,4} To provide this information, AIRTIGHT (unlike many implementation studies) provides a published study protocol, references established frameworks (RE-AIM⁵ and PRECIS-2⁶), and specifies intended outcomes. Despite these efforts, it remains unclear exactly how the intervention was implemented, what the level of adherence to the pre-specified intervention components was, or what the timing of the intervention relative to the evaluation was. As a QII, it is likely that the intervention underwent Plan-Do-Study-Act cycles prior to initiating randomization, and it is also likely that the intervention evolved during the formal study period. More detail on the implementation process would give a clear picture of what was feasible and of what specific intervention components the study best tests. The study results would then better inform us about whether and how to move forward on reducing readmissions.

System leaders and managers charged with allocating resources for a new program such as AIRTIGHT must consider the program's potential capacity for providing equitable, appropriate access across an intended population of users. This is

an understudied issue in the practice of implementation. A starting point would be a flow chart showing how many patients received intervention materials, were contacted by the navigator, accepted or refused engagement, and were contacted after engagement. Review of the flow chart could better identify program capacity, bottlenecks, and patient acceptance. Collecting this information as part of a pragmatic or quality improvement intervention is not always feasible, given that research staff involvement is minimal; methods for improving the feasibility of data collection on sample flow, such as through use of low burden online templates, are needed.

Additional ideas for future investigation can be gained by juxtaposing AIRTIGHT with the two randomized trials cited by the authors as testing similar interventions. One was based in Singapore⁷ and had strongly positive results and one was based in Toronto⁸ and had negative results. The positive Singapore study was carried out in a fragmented system without reliable continuity primary care. We might suppose that a transition-focused intervention would be particularly useful when existing transition management was problematic. The negative Toronto study, like AIRTIGHT, was carried out within a delivery system (the Canadian Health System) known to have (despite some gaps cited by the authors) a strong continuity primary care base and at least basic existing care transition linkages. Addressing remaining readmissions in highly capable systems such as these may require different types of interventions. Additionally, the AIRTIGHT 30-day readmission rate across study groups of 16% may be a minimum; it might or might not be cost effective to reduce readmissions further.⁹

In a delivery system, the selection and enrollment of patients are realistic and potentially expensive parts of what must be done to make a program work. While identifying high-risk patients electronically is intuitively appealing, it produces a heterogeneous patient group. Some patients are optimally cared for without additional intervention, some patients are not willing to participate, and the remainder have indications for a diverse range of clinical interventions. This poses major intervention design challenges. Illustrating these challenges, a lower than expected sample of electronically selected patients in AIRTIGHT experienced the target intervention components. The resulting sample size issues somewhat reduced the strength of the study's conclusions on effects of transition assistance but highlighted the effort required to engage computer-identified high-risk patients. In future work, feasible additions to electronic high-risk patient identification could be accounted for and tested, such as clinician identification and referral followed by electronic screening, electronic identification followed by record review and survey-based screening, or intervention testing in "hot spot" hospitals with unexpectedly high readmission rates.

To improve translation of high-risk patient-related recommendations into practice, we need to rethink patient flow and how it is accounted for in the studies upon which the recommendations are based. In AIRTIGHT, as in both the Singapore and Canadian studies, over 70% of electronically assessed

high-risk patients did not receive the transition assistance program. In the Singapore and Canadian studies, unlike the AIRTIGHT study, randomization occurred after additional eligibility criteria were applied to the electronically selected patients, and after patients agreed to participate. The AIRTIGHT evaluation tests the process of contacting, assessing, and engaging patients as part of its intervention, whereas the other two studies mask the effects of these steps by carrying out the bulk of them as pre-intervention activities. Randomizing only patients whom research staff determine to be eligible and willing is helpful for focusing evaluation on innovative downstream intervention features. The AIRTIGHT approach, however, is a truer reflection of what an implementing delivery system is likely to experience, and more appropriate for testing readmission recommendations. Future recommendations on high-risk patients should include consideration of patient selection and engagement challenges.

Once engaged, AIRTIGHT participants underwent clinical assessment by a physician and other team members. However, no systematic assessment tools are referenced. Geriatric assessment and management approaches, as applied in a wide variety of randomized trials, can reduce hospitalization in heterogeneous elderly populations.¹⁰ These approaches emphasize clinical identification of "root cause"-type problems such as dementia, depression, fall or pressure ulcer risk, and hearing or vision deficits as an important part of clinical treatment or management planning. High-risk patient intervention studies that specify and test linkages between feasible clinical assessments and resulting clinical interventions are needed.

In summary, while rigorous research studies of readmission interventions provide a valuable basis for implementation, we may repeat variations on them for too long without testing their conclusions in delivery systems. In part, we may fear failure, which is the rule, not the exception, in implementing research-based interventions within realistic delivery system management structures and constraints. Thought leaders, however, identify "failure" as the best path to success, and the only way to truly innovate. The ability to experience joy, enthusiasm, and renewed creativity from both positive and negative delivery system studies is a cornerstone of the practice of implementation science and reflects commitment to ongoing healthcare improvement.

Corresponding Author: Lisa V. Rubenstein, MD, MSPH; University of California Los Angeles David Geffen School of Medicine and Fielding School of Public Health RAND Corporation, Santa Monica, CA, USA (e-mail: Lisar@RAND.Org).

REFERENCES

1. McWilliams A, Roberge J, Anderson WE, et al. Aiming to Improve Readmissions Through InteGrated Hospital Transitions (AIRTIGHT): a Pragmatic Randomized Controlled Trial. *J Gen Intern Med*. 2018. <https://doi.org/10.1007/s11606-018-4617-1>.

2. **Danz MS, Rubenstein LV, Hempel S**, et al. Identifying quality improvement intervention evaluations: is consensus achievable? *Quality & safety in health care*. 2010;19(4):279–283.
3. **Davidoff F, Batalden P, Stevens D, Ogrinc G, Mooney SE, group Sd**. Publication guidelines for quality improvement studies in health care: evolution of the SQUIRE project. *BMJ (Clinical research ed)* 2009;338:a3152.
4. **Hempel S, Shekelle PG, Liu JL**, et al. Development of the Quality Improvement Minimum Quality Criteria Set (QI-MQCS): a tool for critical appraisal of quality improvement intervention publications. *BMJ Qual Saf*. 2015;24(12):796–804.
5. **Glasgow RE, Klesges LM, Dzewaltowski DA, Estabrooks PA, Vogt TM**. Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. *Health education research*. 2006;21(5):688–694.
6. **Loudon K, Zwarenstein M, Sullivan FM**, et al. The PRECIS-2 tool has good interrater reliability and modest discriminant validity. *J Clin Epidemiol*. 2017;88:113–121.
7. **Low LL, Tan SY, Ng MJ**, et al. Applying the Integrated Practice Unit Concept to a Modified Virtual Ward Model of Care for Patients at Highest Risk of Readmission: a Randomized Controlled Trial. *PLoS One*. 2017;12(1):e0168757.
8. **Dhalla IA, O'Brien T, Morra D**, et al. Effect of a postdischarge virtual ward on readmission or death for high-risk patients: a randomized clinical trial. *JAMA*. 2014;312(13):1305–1312.
9. **Reuben DB, Tinetti ME**. The hospital-dependent patient. *The New England journal of medicine*. 2014;370(8):694–697.
10. **Stuck AE, Iliffe S**. Comprehensive geriatric assessment for older adults. *BMJ (Clinical research ed)* 2011;343:d6799.