

## TO THE EDITORS

We are delighted to see this study by Javitt et al<sup>1</sup> with its disclosures of financial interests, detailed description of methods, and use of randomized design in a real-world setting. Indeed, the study presents compelling evidence regarding the clinical value of the algorithms that generate the alerts (termed “care considerations”). However, one concern we have regards the change in the intervention group’s per member per month (PMPM) costs, relative to the reference group’s PMPM, among the subgroup of the intervention population that did not receive any alerts. We estimate this change to be \$5.00 PMPM (based on data in Tables 1 and 3).

Group	N	Impact (PMPM)
Alerts	961	\$68.08 (supplied)
Non-Alerts	18 778	\$5.00 (calculated)
Total	19 739	\$8.07 (supplied)

We have considered 2 explanations for this \$5.00 drop in PMPM in the intervention group members who did not receive the alerts. First, the randomization scheme did not produce equivalent reference groups at baseline on factors other than age and sex (which were found, overall, to be not statistically significant), and nonequivalence on clinically significant risk factors was projected to the follow-up period. The second possible explanation is that influences besides the alerts, post-baseline, were directed, inadvertently or not, to the intervention subgroup that did not trigger alerts.

To explore the first, we suggest the authors examine in more detail baseline characteristics beside the age and sex characteristics listed in Table 1. Risk factors available for use in administrative data include diagnoses, comorbidities, procedures, place of service encounters, and costs. If the groups were indeed equivalent at baseline, we suspect that something happened to make them nonequivalent in the follow-up period.

To explore this possibility, we would suggest these options: (a) Conduct qualitative research into interventions that may have been directed to the intervention group only and report the findings; and/or (b) employ a matched case-control design in which reference group patients are matched (direct matching or matching on the propensity score<sup>2</sup>) to a representative sample of intervention patients who did not trigger alerts.

Depending on the results of these additional inquiries, we would also suggest a recalculation of

return on investment (ROI) in which the numerator value is based on those PMPM changes that can be attributed to the intervention(s). If the alerts alone were considered the intervention—and the intervention and reference groups were equivalent at baseline on important factors besides age and sex—the total cost of the program would be based on everyone in the intervention population (n = 19 739 or \$19 739), while the total PMPM impact would only be based on those who received “care considerations,” or alerts (n = 961 or \$65 425). Thus, the 18 778 patients (or \$93 868) who did not receive alerts would be excluded from the analyses and the ROI would be a positive value of 3.31 (\$65 425/\$19 739), rather than 8.07 (\$159 229/\$19 739) as reported by Javitt et al.

Again, we compliment the authors for publishing this study and providing sufficient detail to allow others to comment. We are acutely aware of the great challenges of performing rigorous studies in real-world situations and appreciate such important work. We do not seek “scientific proof,” but rather the following of well-accepted evaluation principles such as those outlined by the Population Health Impact Institute (with which we are both associated).<sup>3</sup> The best way to advance the cause of disease management and other defined population health management programs is to follow the tenets of science: publish, with disclosure of methods in sufficient detail to allow others to replicate findings, discuss methods and results, correct errors, and learn together.

Thomas Wilson, PhD, DrPH  
Wilson Research, LLC  
Loveland, Ohio

Ariel Linden, DrPH, MS  
Linden Consulting Group  
Hillsboro, Ore

## REFERENCES

1. Javitt JC, Steinberg G, Locke T, et al. Using a claims data-based sentinel system to improve compliance with clinical guidelines: results of a randomized prospective study. *Am J Manag Care.* 2005;11:93-102.
2. Linden A, Adams J, Roberts N. Using propensity scores to construct comparable control groups for disease management program evaluation. *Dis Manag Health Outcomes.* 2005;13:107-127.
3. Population Health Impact Institute. The five evaluation principles. Available at: <http://www.phiinstitute.org/evaluation.html>. Accessed April 10, 2005.

(Response on page 394)

*(Continued from page 384)*

**IN REPLY**

The correspondents have noted that the data suggest a possible difference between the intervention group and the control group, favoring the intervention group. In contrast to the \$68 PMPM difference between cases and controls who trigger the intervention, the difference between those who do not trigger the intervention is \$3.13 PMPM ( $P = .045$ ), which was not statistically significant in our judgment and was so noted in the text. Our belief is that this difference simply falls within the random variation of healthcare utilization and is not attributable to any baseline difference between the 2 randomly selected groups.

We ask readers to focus their attention on the 20-fold larger difference among those triggering the intervention. If the correspondents wish to suggest that our intervention merely saved 5 times its cost, rather than 8 times its cost, we will not engage in theological argument but do hope that they and the readership don't

miss the main point of the study. The point is that a randomized, prospective trial, conducted in a representative urban population, demonstrates that the use of a claims data-based decision support system achieves major savings in morbidity, cost, and ultimately lives. The observed difference is attributable to those subjects who triggered the intervention. Moreover, the subgroup analysis for those triggering recommendations related to the HOPE trial showed that the difference in hospitalizations was totally accounted for by cardiac and pulmonary admissions as would be expected from the primary endpoints of the HOPE trial. In the near future, we will be presenting the results of a confirmatory trial, conducted in 4 other geographic areas that point to the same conclusion.

Jonathan C. Javitt, MD, MPH  
*Potomac Institute for Health Studies*  
*Washington, DC*