**Supplemental Digital Content 3: Generalizability of Results**

A key issue is the definition of the population of interest.  Because "early adopting" HDHP employees differ from those in more traditional plans, the data might not support inference about the effects of an HDHP on the entire employed population (population average treatment effect, ATE).  Instead we matched each case in the HDHP sample to a control case with similar characteristics. This enabled us to estimate effects for a population of employees with characteristics like those in HDHPs (average treatment effect for the treated, ATT).

We performed a member-level and not an employer-level analysis and, although employers that selected HDHPs might be somewhat unusual (in that they were early adopters or experiencing greater cost pressures), it does not follow that employees and dependents would be especially unrepresentative or unusual. Furthermore, we believe that our propensity score matching approach was the most rigorous method available to maximize internal validity with likely minimal detriment to external validity compared with other approaches. (Ideally we would have restricted analyses to the strata of members with middle propensity scores to maximize external validity [e.g., the middle tertile] but our sample size was insufficient for this approach.)

In summary, we believe that our results are generalizable to persons enrolled through small employers that select into HDHPs that do not subject outpatient visits and preventive labs to the deductible. This is a common design in the employer-based insurance market and will likely be a common offering in health insurance exchanges (California, Oregon, and Vermont offer such arrangements). Enrollment in such plans is therefore likely to become increasingly common after the 2014 Affordable Care Act coverage provisions are enacted. However, a careful and measured discussion of generalizability is prudent, which we included in the limitations section of our study.