Retrospective Data Review and Propensity Scoring: Religion (Believing) or Science (Proving) and the Appropriate Application of Statistics

To the Editor:

A recent article presented for continuing medical education credit studied perioperative statin therapy in elective aortic surgery during 2001–2009, a period of significant advances in aortic surgery, statin use, and perioperative management.[1] Although propensity scoring (PS) was used to balance some variables related to the choice of statin exposure *versus* outcomes, validity demands that *all relevant parameters* be subjected to validate the methodology.[2] Although the authors specifically stated: "The frequency of treatment (statins) according to the year of surgery increased significantly with time" and "chronic statin therapy is used in association with other cardiovascular medications," the year of surgery as a variable was not considered in the PS adjustment relative patient outcome.

How did advances and trends toward endograft treatments (annual rates of endograft *vs.* open abdominal aortic aneurysm, did any patients receive endograft or were converted to open procedures?), variable aneurysm type and diameter, abdominal *versus* retroperitoneal approach, suprarenal cross-clamping incidence and time, surgeon experience, therapy options, blood salvage techniques, American College of Cardiology/American Heart Association guideline introduction and updates including statin and *repeated* revision of β-blockade recommendations (as well as Perioperative Ischemic Evaluation (POISE) study results affecting β-blockade utilization), affect any care delivered and outcomes in 2001 (before American Heart Association/American College of Cardiology guidelines) *versus* 2006 *versus* 2009! The duration and indication of preoperative statin use, time from aortic disease diagnosis to surgery, smoker *versus* nonsmoker (baseline carbon monoxide-hemoglobin concentration), preoperative lipid and C-reactive protein levels, as well as the frequency and quality of primary medical care before surgery, would have been additional important data to consider in applying PS. Although "all patients were screened in accordance of American College of Cardiology/American Heart Association guidelines," the findings of Polderman's very astute Erasmus group's care raised serious questions concerning just how these guidelines are/were actually implemented, or conversely, specifically resulted in any observed increased incidence of statin-treated patients in later years.[3,4] The year of surgery is specifically worthy of inclusion in the study's PS analysis, given the author's own clear observation. Many other factors not analyzed would be expected to have greater effect than statin therapy itself. Are findings valid without these multiple listed considerations, yet alone those differences specifically noted but not considered by the authors?

The presence of statins and other drug therapies may specifically indicate superior presurgical care states, medical and patient education advances over a period of years, or simply modern patient compliance and concern. Active patient-directed pharmaceutical advertisements are commonplace in the United States and resulted in high statin usage; is this also typical of France and Europe, leading to the very high usage today *versus* 2001? The 21 patient characteristics included 12 of 21 parameters, which demonstrated statistically significant differences between the groups, "equalized" by PS. What *are* the really important available parameters? Was PS meticulously used and how does this reflect on results, conclusions, reader's interpretations, and continuing medical education value/emphasis of the paper? "First of all, as propensity score can only remove overt (known) bias but unlike randomization it cannot be expected to remove hidden (unmeasured) bias, results from its application should be considered with caution. Interpretation will depend on the quality and amount of information about the efficacy of the treatments under evaluation."[2] It may be time to introduce ongoing education regarding statistical analysis as a necessary component of the education section of this journal and a component of all continuing medical education review articles, given the important ongoing changes in statistics and importance to medical decisions and ultimately, patient care. PS should/can ultimately/only lead to randomized trials to confirm a PS *suggested observation*, whenever possible.[5]

Paul Martin Kempen, M.D., Ph.D., Cleveland Clinic, Cleveland, Ohio. kempenp@ccf.org

References

1. Le Manach Y, Ibanez Esteves C, Bertrand M, Goarin JP, Fléron MH, Coriat P, Koskas F, Riou B, Landais P: Impact of preoperative statin therapy on adverse postoperative outcomes in patients undergoing vascular surgery. ANESTHESIOLOGY 2011; 114:98–104
2. Cavuto S, Bravi F, Grassi M, Apolone G: Propensity score for the analysis of observational data: An introduction and an illustrative example. Drug Dev Res 2006; 67:208–16
3. Hoeks SE, Scholte op Reimer WJ, Lenzen MJ, van Urk H, Jörning PJ, Boersma E, Simoons ML, Bax JJ, Poldermans D: Guidelines for cardiac management in noncardiac surgery are poorly implemented in clinical practice: Results from a peripheral vascular survey in the Netherlands. ANESTHESIOLOGY 2007; 107:537–44
4. Fleisher LA, Beckman JA, Brown KA, Calkins H, Chaikof E, Fleischmann KE, Freeman WK, Froehlich JB, Kasper EK, Kersten JR, Riegel B, Robb JF, ACC/AHA TASK FORCE MEMBERS, Smith SC Jr, Jacobs AK, Adams CD, Anderson JL, Antman EM, Buller CE, Creager MA, Ettinger SM, Faxon DP, Fuster V, Halperin JL, Hiratzka LF, Hunt SA, Lytle BW, Md RN, Ornato JP, Page RL, Riegel B, Tarkington LG, Yancy CW: ACC/AHA 2007 Guidelines on Perioperative Cardiovascular Evaluation and Care for Noncardiac Surgery: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Com-

 Anesthesiology 2011; 115:656–62

mittee to Revise the 2002 Guidelines on Perioperative Cardiovascular Evaluation for Noncardiac Surgery): Developed in collaboration with the American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Rhythm Society, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, and Society for Vascular Surgery. Circulation 2007; 116:1971–96

5. Nuttall GA, Houle TT: Liars, damn liars, and propensity scores. ANESTHESIOLOGY 2008; 108:3–4

(Accepted for publication May 5, 2011.)

In Reply:

Regarding the number of letters Dr. Kempen produced in the past years, his comment relative to our study about perioperative use of statins[1] might be useful for understanding perioperative effect of chronic cardiac treatment.

The first part of the comment was related to the inclusion of variables in the propensity score. We did not include the variable "Year of surgery" in the propensity score because it did not predict the presence of statins. A better understanding of the method would have allowed consideration that patients' medical history might have changed during the year in question, and that the expected effect of this covariable on the probability of being treated by statins was implicitly characterized by other variables that were already included in the propensity score.

As we have shown, the adjustment of the propensity score provided a very good balance between statin users and control subjects. The inclusion of the year of surgery would not have improved the propensity score, as well as the other variables mentioned in the comment. The second part of the comment was related to the surgical characteristics. The rate of endograft procedures remained constant, and as low as 7% during the study period. This low rate, associated with the high experience of the surgeons, probably explains why we did not observe any conversion to open procedures. Endograft treatment was not associated with statin use. Furthermore, although probably not outlined clearly enough, these patients were not included in the analysis, because the rate of postoperative major adverse events is recognized as being lower.

The difference between abdominal and lumbar approach was not associated with the preoperative use of statins, when the other variables were taken into account. In fact, it might be obvious for most of us that some surgical characteristics are predictors of postoperative adverse events. However, the question is to know whether these events are associated with the use of statins. This is the aim of propensity score, *i.e.*, to estimate the probability of receiving the treatment of interest given the other covariates.

In randomized controlled trials, treatment allocation is at random, meaning that the probability of receiving the experimental treatment is 0.5 for each patient. Conversely, for propensity score analysis, the probability of receiving the experimental treatment, given the other covariates, is identical for each patient. However, because some unknown and underlying covariables might operate, a propensity score analysis should not conclude about causality. Consequently, randomized controlled trials will always remain the gold standard.

Finally, the comment of Dr. Kempen underlines the need for education about modern statistical methods. We fully agree upon the role of propensity score methodology in exploring drug effects as well as in the need of review articles for emergent and complex statistical methodology.

Yannick Le Manach, M.D.,* Bruno Riou, M.D., Ph.D., Paul Landais, M.D., Ph.D. *Centre Hospitalo-Universitaire Pitié-Salpêtrière, Paris, France. yannick.le-manach@psl.aphp.fr

Reference

1. Le Manach Y, Ibanez Esteves C, Bertrand M, Goarin JP, Fléron MH, Coriat P, Koskas F, Riou B, Landais P: Impact of preoperative statin therapy on adverse postoperative outcomes in patients undergoing vascular surgery. ANESTHESIOLOGY 2011; 114:98–104

(Accepted for publication May 5, 2011.)

Relevance and Value of a Morphine Immunoassay as a Diagnostic Aid for Neuromuscular Blocking Drug-induced Anaphylaxis

To the Editor:

Neuromuscular blocking drugs (NMBDs) are the major cause of episodes of immediate hypersensitivity during the perioperative period.[1] The detection of liberated mast cell tryptase[2] is used to confirm an anaphylactic event; skin prick testing with all currently used unconjugated NMBDs, followed by intradermal tests if necessary, is central to identifying the culprit, as well as cross-reacting drugs. Identification of the implicated NMBD(s) is important for providing safe anesthesia for the patient in the future. Although skin tests remain the diagnostic tool of choice, their sensitivity and specificity have at times been questioned. Serum immunoglobulin E (IgE) antibody assays are often used in cases involving skin test-negative or equivocal reactors or when skin tests are unreliable or unavailable, making these tests useful adjuncts to skin and tryptase tests. They are also valuable when applied to sera taken at the time of the reaction, to preoperative serum samples, and serum taken before or after death.[3] The IgE assay in inhibition form can sometimes help in identifying and providing immunochemical insights into cross-reacting drugs.[3,4] After extensive testing over many years on hundreds of serum samples, a morphine-solid phase was shown to be superior to NMBD-solid phases and other selected solid phases for the detection of NMBD-reactive IgE antibodies, and its use, along with the tryptase and skin tests, was strongly advocated as the best diagnostic combination of