Appropriate interpretation of the observed NOACs treatment effect size in real-world study

Yang Wang, MS[a], Tao Chen, PhD[b], James Dodd, Msc[b],Wei Li, PhD[a]

[a]*Medical Research & Biometrics Center, National Center for Cardiovascular Diseases, Fuwai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China*

[b]*Tropical Clinical Trials Unit, Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, L3 5QA, UK*

Correspondence author: Yang Wang

Medical Research & Biometrics Center, National Center for Cardiovascular Diseases, Fuwai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Room 101-106, BLOCK A Feng Cun Xi Li, Yongding Zhen Mentougou District Beijing, 102300, China

Phone: (86 10) 6086 6499

Mobile: (86) 138 1022 7427

E-mail: wangyang@mrbc-nccd.com

Funding Sources: None.

Conflict of Intrest: None declared. All authors have participated in the preparation of

this manuscript.

Article type: Letter to the Editor

Key words: Effect size, Real-world study

Running head: Effect size in real-world study

To the Editor:

Dr. Graham et al reported their research on the comparative effectiveness of warfarin and Nonvitamin K antagonist oral anticoagulants (NOACs) amongst patients with nonvalvular atrial fibrillation. This real-world study was based on claims data (US Medicare).[1] The authors concluded that dabigatran and apixaban were associated with a more favorable benefit−harm profile than rivaroxaban in standard-dose NOAC users.

We also note Dr. Graham's previous work in JAMA Internal Medicine 2016.[2] This study enrolled the Medicare nonvalvular atrial fibrillation patients who initiated dabigatran or rivaroxaban for stroke prevention from November 2011 to June 2014. Data for the recent paper were taken from October 2010 to September 2015, and it seems the two studies may have some overlap, because they have similar inclusion/exclusion criteria and results for the comparison between dabigatran and rivaroxaban. Can the authors to clarify the differences between current and 2016 paper? For example, have identical covariates had been used in propensity score (PS) models? Did the PS model selection perform in the same dataset involving outcome data in it?[3] The authors mention both 1:1 PS matching and stabilized inverse probability weighting (IPW) in the current analysis; estimated treatment effects can vary greatly between different PS methods using the same data.[4]

The observed effect size was relatively small in the recent report, with weighted event rates of intracranial hemorrhage were 3.3 and 4.8 per 1,000 person-years for dabigatran and apixaban, respectively. The between group difference was statistically significant (95% confidence interval of hazard ratio excluded 1.0). We performed a power analysis with the assumed between group effect size (0.15% ≈ (4.8-3.3)/1,000) under different sample size scenarios (see Table 1). The current study (N=86,198 for dabigatran and N=73,039 for apixaban) seems over-powered to show a statistically significant between group difference which may not be clinical relevant. The dramatic decrease in the number at risk among the Kaplan-Meier plots should also be noted as it indicates that majority of the patients contributed very short-term follow-up information.

In conclusion, real-world data can help us to compare different NOACs for effectiveness or safety, but the transparency of the analysis process and validity of the statistical method are paramount. The effect size for a specific intervention from an observational study should be interpreted with caution unless proven by randomized control trial.

Acknowledgments: None.

References:

1. Graham DJ, Baro E, Zhang R et al. Comparative Stroke, Bleeding, and Mortality

Risks in Older Medicare Patients Treated with Oral Anticoagulants for Nonvalvular Atrial Fibrillation. *Am J Med*. 2019;132(5):596-604.e11

2. Graham DJ, Reichman ME, Wernecke M et al. Stroke, Bleeding, and Mortality Risks in Elderly Medicare Beneficiaries Treated With Dabigatran or Rivaroxaban for Nonvalvular Atrial Fibrillation. *Jama Intern Med*. 2016;176(11):1662-1671
3. Yue LQ. Regulatory considerations in the design of comparative observational studies using propensity scores. *J Biopharm Stat*. 2012;22(6):1272-9
4. Kurth T, Walker AM, Glynn RJ et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *Am J Epidemiol*. 2006;163(3):262-70