Dear Editor,

We would like to highlight two fundamental problems in aspects of the design and statistical analysis utilised by Harris et al. in their paper which was published in the JME 2019;22(7):691-697. We think the problems which we have identified with the matching of the myasthenia-gravis (MG) patients to non-MG patients and the statistical analysis applied to the resulting retrospective matched cohort data were capable of affecting the validity of some of the reported study findings.

**Study design-matched cohort of refractory MG and non-MG patients**

The non-MG patients were randomly matched to each MG patient at a ratio of 4:1 on age at index date, sex and general practice as controls, having been assigned the same index date as that of the MG patient (i.e. date of first diagnosis of MG) and with the additional requirement of having at least 12 months of data available prior to the index date. The follow-up period was defined as "the day after the index date until the earliest of the following: (a) the patient transferred out of the practice, (b) the last date of data collection or (c) the study end date", which according to the reported results ranged over several years in duration. In other words, the authors had assumed a time-to-event framework for their matched cohort, without due regard to its implication on the main endpoints of interest to the study, which they had summarised as healthcare resource utilisation.

According to the literature, the validity of the matched cohort design and analysis of the resulting data assume as applied by the authors would hold only if there is no loss to follow-up. Consequently, if some individuals are lost during follow-up, then the cohorts are only matched at the starting date (i.e. at index date). Indeed, the longer the duration of follow-up, the more impactful would be the problem of increasing loss of comparability between the two groups on the study findings. This is likely in the case of differential follow-up periods, and/or, differential attrition between cohorts.

**Statistical analysis of the matched cohort**

It is, of course, impossible to control losses to follow-up during the study period and this poses a major problem which was not addressed in the statistical analysis, making comparability of the MG patients and their controls difficult to justify. The paper reported the median length of follow-up (interquartile range [IQR]) as 74.28 (43.45–111.81) months for the refractory MG cohort and 63.19 (28.57–97.73) months for the non-MG control cohort (p=:.0704). The median difference in follow-up of more than 11 months between the two groups could not be justifiably ignored as irrelevant, especially for such study endpoints that involved count data over the follow-up period. Indeed, none of the endpoints of interest was about time-to-event (i.e. risk assessment).

By failing to provide any justification for inadvertently assuming that the matched groups would remain comparable over the years of follow-up periods, the authors were wrong on two counts: (1) by their ignoring the matching variables in the statistical analysis, and more importantly, (2) by estimating count endpoints over unequal periods of follow-up between the two groups. The person-years approach could not adequately account for the likelihood of non-comparability. Moreover, their analysis does not provide information on the pattern of utilisation over time.

The mistakes we have highlighted are not uncommon among burden of disease as well as healthcare resource utilisation studies in real world settings. These studies usually involve the comparison of patients with a particular disease of interest with those free of the disease (as controls) and the matched cohort design is very popular. The impact of the problems described above might be minimized if the follow-up duration is restricted to a short time-frame – the use of a maximum of one-year follow-up period is common. The outcomes of interest for these studies are often based on count data, which rely on equitable duration of assessment between the two groups for comparison.

A pragmatic solution, which might minimize these problems to an acceptable level (i.e. without significantly impacting the study results) would be to ensure that only non-MG patients with similar follow-up durations as that of the MG patient, are eligible for consideration as potential controls to the MG patient. In practice, we match each MG patient to eligible non-MG candidates additionally (i.e. on duration of follow-up), such that the follow-up periods for the matched patients- the MG patient and control(s) are within a short interval of each other, such as a month or a little more but no more than a year.

Like the authors, we have assumed that apart from the matching variables, there are no other significant sources of bias in terms of measured and unmeasured confounding. We acknowledge that the resulting controls would be a limited subset of the general population of non-MG patients (i.e. only those with similar enrolment record as the MG
population in the database). However, the approach could significantly reduce the problem of comparability of the two groups, which is critical to the validity of the study results. This will restrict the source of the comparability problem to the increasing loss of the matching effect due to the increasing duration of follow-up, which can be addressed by accommodating the matching variables in the statistical analysis.

Transparency

Declaration of funding

No funding was received in connection with the letter

Declaration of financial/other relationships

No potential conflict of interest was reported by the author.

Author contributions

Both authors were involved in the drafting of the letter.

Acknowledgements

None to acknowledge.

References

[1] de Graaf MA, Jager KJ, Zoccali C, et al. Matching, an appealing method to avoid confounding? Nephron Clin Pract. 2011;118(4):c315–c318.

[2] Rothman K, Greenland S, Lash T. Modern epidemiology. 3rd ed. Philadelphia (PA): Lippincott Williams & Wilkins, 2008.

[3] Jewell N. Statistics for epidemiology. London (UK): CRC Press, 2004.

[4] Sjölander A, Greenland S. Ignoring the matching variables in cohort studies – when is it valid and why? Stat Med. 2013;32(27):4696–4708.

[5] Walker AM, Jick H, Hunter JR, et al. Vasectomy and non-fatal myocardial infarction. Lancet. 1981;1(8210):13–15.

[6] Walker AM. Efficient assessment of confounder effects in matched follow-up studies. Appl Stat. 1982;31(3):293–297.

[7] Walker AM, Jick H, Hunter JR, et al. Vasectomy and nonfatal myocardial infarction: continued observation indicates no elevation of risk. J Urol. 1983;130(5):936–937.

[8] Cummings P, McKnight B, Greenland S. Matched cohort methods for injury research. Epidemiol Rev. 2003;25:43–50.

[9] Soriano JB, Visick GT, Muellerova H, et al. Patterns of comorbidities in newly diagnosed COPD and asthma in the primary care. Chest. 2005;128(4):2099–2107.

[10] Kiri VA. Description of an incidence-based model for assessing comorbidity patterns in disease natural history. BMJ Open. 2016;6(7):e012105.