Reviewer’s report

Title: Heavy episodic drinking and soccer practice among adolescents in Brazil: the contextual aspects that involves this relationship

Version: 1 Date: 4 December 2012

Reviewer: Lionel Riou França

Reviewer’s report:

The study provides an interesting analysis of the influence of the type of sports practice on heavy episodic drinking. It originates from a large, representative sample of urban 14-18 year old students. The main hypothesis is clearly identified and the results are extensively discussed.

Discretionary Revisions

1/ Heavy episodic drinking seems to have been measured in the same way for males and females. Some authors believe that the criteria of 5 drinks for males and 4 drinks for females on one occasion is more accurate (e.g. Wechsler, Dowdall, Davenport, & Rimm, 1995). What is the rationale for the authors’ choice?

2/ While the possibility of over-reporting has been considered, the possibility of under-reporting is not. the authors should discuss the possibility of under-reporting of own substance use in the context of self-reported use questionnaires.

3/ The authors clearly state that their sample is representative of Brazilian state capitals. Do they feel the results would have been different in more rural areas?

4/ Some background information about the proportion of 14-18 year old Brazilians that don't go to school would be appreciable to have an idea about the generalizability of the results.

5/ In table 2, the adjusted and crude OR for soccer players vs. sedentary students are the same, including the confidence intervals. Is this the result of a typo? If the ORs are genuinely the same, this could be specified in the text.

6/ Was the interaction between sex and sport practice tested? In other words, is the association between sport practice or soccer practice and HED the same for males and for females?

Minor Essential Revisions

1/ The authors should consider the submission of their article to an English speaking corrector. As English is not my mother language, I cannot identify all errors, but some errors/typos are present (e.g. “for some analysis” instead of “analyses”).
2/ In section 2.3, the authors should use sentences to present the collected variables instead of a bullet-point type of presentation (e.g. “Sociodemographic variables: were obtained by (…)”).

3/ In section 2.4, I understand the use of LRM1 and LRM2 for the models, but what does the “OR” term in the “crude model (OR)” sentence stand for?

4/ In section 3.1, the authors refer to the frequency of HED in the last month but report the figure for the “6 or more times” modality. Are they referring to HED 6 or more times per month, or to HED at least once in the last month (which would be of 21.0%)?

5/ In tables 1 to 3, the meaning of “wgt” (I assumed it meant weighted) should be specified.

6/ The authors state in the discussion that “Our study indicated that the induced sociability could be a more important factor than the sport practiced in relation to alcohol use”. This seems to be contradictory with the finding that soccer playing is more associated to HED than other team sports. Please clarify. Furthermore, the respective influences of socialization and sport practice are hard to infer since we do not have the data about students that practice sports that are not team sports.

Major Compulsory Revisions

1/ In order to appreciate the sample’s representativeness and the external validity of the results, more information is needed about the sampling procedures and inclusion criteria. The abstract should state that all states are sampled (as this information is more relevant than the number of states sampled). Furthermore, a flow chart of the students included in the analysis would improve understanding. The authors collected 19,230 questionnaires and only analyzed 9,886. Why were the others excluded? What is the proportion of non-sedentary, non team sports players and why are these students excluded? What is the proportion of students not reporting their HED status?

2/ In order to rule out possible biases in the association between HED and sport practice, I would be interested to see:

(1) a comparison in the rate of non-response for HED between soccer players and the others. If non-response is important and different in the compared groups, the estimates might be biased.

(2) a comparison in the rate of students reporting fictitious drug use between soccer players and the others. Although fictitious drug use reporting is uncommon, a difference in its reporting could indicate that over-reporting might be more likely in some groups.

Furthermore, 20.5% of the students were absent in the day of the survey. Could both alcohol use and sport practice differ among absentees? This issue should be discussed in the limitations section.
3/ Table 1 should present a test of association between the groups for all variables presented in order for the reader to quickly identify which characteristics are unbalanced.

4/ The authors state that “the individual should be influenced by normative attitudes of the group, and this bond appears to be more intense in soccer”. On what evidence is this affirmation based? If it is not supported by the data or other studies, the affirmation should be toned down.

5/ Multiple comparisons and tests have been performed in this study, and there has been no adjustment for the inflation of type I error. The authors should acknowledge the exploratory nature of their analyzes in the limitations section.

6/ The authors correctly remind that cross-sectional designs don't allow for the detection of which, of HED or sport practice, occurred first. However, even in longitudinal studies, knowing that X occurs before Y does not guarantee that X caused Y. The authors should also acknowledge that causality cannot be assumed for other reasons. There could be unmeasured factors that bias the association between sport practice and HED. While adjusted analyzes have been presented, the authors only adjust for a small number of variables (age, gender and type of school) and other confounders could be responsible for the reported associations. For instance, the socioeconomic status of soccer players could be different. The possibility of unmeasured confounders should be stated in the limitations section.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.