Reviewer's report

Title: Tanzanian malignant lymphomas: WHO classification, presentation, ploidy, proliferation and HIV/EBV association

Version: 1 Date: 1 December 2009

Reviewer: Pietro Bulian

Reviewer's report:

Major Compulsory Revisions

1) authors should correct many inconsistencies in numbers and statistics (I was not able to reproduce p values) in the following sections, and modify discussion accordingly:

a. page 10 and table 2: the percentages of cases with extranodal disease are incorrect as 54/106 = 50.9% not 49.5% as reported, 9/43=20.9% not 8.3%, moreover I was unable to replicate the statistical figures reported in table 2, namely the comparison between 50.9% and 20.9% (extranodal cases in children vs elderly) gives a p of 0.00089 by fisher test and 0.00077 by chi-square test, but not <0.0001 as reported. The authors probably have calculated the proportions of children/elderly among total extranodal cases (54/109 and 9/109), not the proportion of extranodal cases among total children or elderly patients. I was also unable to reproduce the reported p=0.124 in the comparison of extranodal cases between female and male patients, either by the chi-square test or by fisher test (p=0.31). Please check thoroughly the numbers and the statistical tests.

b. page 11 and table 4: on page 11 the authors state "as expected most extranodal cases were BL (15/17, 88.2%), but these figures are the BL cases with extranodal presentation, the extranodal cases with BL histology should be 15/28 BCL or 15/35NHL or 15/40 ML, according to which denominator the authors might wish to use. The percentage of TCL cases with extranodal presentation could be more consistently compared with BCL cases, not to DLBCL as is in text. Also, please check the numbers and statistics: the fisher test with the numbers provided gives a p value of 0.0949, chi-square test p=0.062, contrary to the p<0.001 reported by authors.

c. page 12 last line: the authors report 12 tetraploid/multiploid cases, a value inconsistent with the numbers reported in page 13 (13 cases).

d. page 13 in line 1 the authors report 9 hyperdiploid cases, a value inconsistent with line 2 (8 cases); moreover the p value reported by authors (p=0.011) seem to be coming from a fisher test on 7/1 vs 1/7 or 87.5% DLBCL vs 12.5% TCL, which is incorrect as it do not take into account the different number of DLBCL and TCL cases, the correct comparison is 7/10 vs 1/5, which is not significant (p=0.369).

e. page 13 line 7: a p value for the comparison is lacking and actually the fisher test gives a p=0.179 (8/9 vs 5/1): the difference is not significant and this should
be declared. The number of tetraploid DLBCL cases (8) is inconsistent with the value reported above in line 5 (7).

f. page 14 line 10 and table 6: the EBER ISH reactivity is unavailable in 2 cases (as reported in table 6), then these cases should be omitted from statistics, the percentage of ABC type DLBCL case showing positivity must be 5/13=38.5% not 33.3% (5/15) as reported in text on page 14. Moreover the p is totally incorrect either with the erroneous percentage (5/15) or the correct (5/13), actually the fisher test on the odd-ratio of 5/8 vs 5/7 gives a p=1 and the chisquare test a p=0.87.

2) authors should avoid the use of any unsupported sentence "not significant" when referring to comparison or tests for which they do not provide numbers or tables and p values. Either the authors provide these numbers or avoid these sentences.

Minor Essential Revisions

1) More details should be provided regarding the context of MNH hospital: is this a reference center collecting most of Tanzanian ML or is it a local area hospital? How much population is served by this hospital? Make a statement explaining if the reported data are to be considered an unbiased representation of whole Tanzanian ML or of ML in a particular geographic area in Tanzania. Are all cases of suspected lymphoma in Tanzania referred to hospitals or a predominance of those highly symptomatic and/or aggressive and/or in young people is to be expected?

2) There are inconsistencies between the cases subjected to hematopathology and IHC (150 in abstract) and the 158 cases of ML with hematopathology and IHC declared in text (page 10). The authors should rephrase the abstract focusing on the number o cases actually confirmed ML (158). It should be stated more clearly (as is done on page 10) that 158 out of 174 biopsies were confirmed to be ML upon revision according to WHO criteria and IHC and that the 174 revised biopsies were a subset of the 336 cases originally classified as ML in Tanzania according to WF classification.

3) the authors report that 109 out of 281 cases with data on clinical presentation had extranodal disease: please detail by which means was the clinical presentation and the extension of disease assessed (e.g., by physical exam only or with Rx, CT scans, ecography...). Please detail also if uniform staging modalities and criteria were applied throughout the considered period (1996-2006) or not.

4) in Abstract-Background it is stated that "WHO has not yet been used in Tanzania": please specify that WF classification is still used (as reported in text).

5) Since this is a retrospective study on archival tissues, please specify if WHO 2001 or WHO 2008 was used, in abstract and text.

6) It is difficult to understand which are the denominators of the percentages
presented in the last column of table 4, since there is not a common denominator and the last row is misleading: the authors should rewrite this table making extra efforts to clarify the various denominators, e.g. total BCL and TCL percentages should be indented (not in line with the percentage of the BCL subgroups), total NHL and HD should be double indented (not in line with BCL subgroups and neither with TCL and BCL). Moreover please thoroughly check the percentage as TCL are 22 out of 134 NHL which is 16.4% not 13.7. Or, alternatively, recalculate all the percentages using the common denominator 158.

7) at page 11 when comparing extranodal cases between TCL and DLBCL the reference should be at table 4 not table 3 as in text.

8) at page 12, at the end of the section on DLBCL and in the section "cytomorphometry results", the authors use the term "apparently not significant" or "appeared to be not significant": a) it would be less ambiguous to say directly "not significant", b) the authors should provide p values.

9) page 14 line 16-18: please provide the actual number of aneuploid cases showing high Ki-67.

Discretionary Revisions

1) The definition "body cavity-based lymphoma" is redundant and may be omitted.

2) page 13, line 3-4: the percentage of triploid DLBCL (66.6%) and BL (33.3%) could be substituted by crude numbers since we know that they are only 3 cases (2 DLBCL and 1 BL).

3) page 13 line 4-5: "Furthermore, the majority of the tetraploid ML...". "the majority of" should be omitted. The term "Furthermore" is repeated and could be omitted.

4) please specify if the cases studied with FC (60), ISH (37) and HIV-elisa (35) are among the 150 (or 158? see above) cases also studied by histopathology and IHC.

Level of interest: An article of limited interest

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