

## ESPS PEER-REVIEW REPORT

**Name of journal:** World Journal of Gastroenterology

**ESPS manuscript NO:** 33150

**Title:** High expression of anti-apoptotic protein Bcl-2 is a good prognostic factor in colorectal cancer: result of a meta-analysis

**Reviewer's code:** 01047751

**Reviewer's country:** United Kingdom

**Science editor:** Yuan Qi

**Date sent for review:** 2017-02-08 13:45

**Date reviewed:** 2017-02-20 09:01

CLASSIFICATION	LANGUAGE EVALUATION	SCIENTIFIC MISCONDUCT	CONCLUSION
<input type="checkbox"/> Grade A: Excellent	<input type="checkbox"/> Grade A: Priority publishing	Google Search:	<input type="checkbox"/> Accept
<input checked="" type="checkbox"/> Grade B: Very good	<input checked="" type="checkbox"/> Grade B: Minor language polishing	<input type="checkbox"/> The same title	<input type="checkbox"/> High priority for publication
<input type="checkbox"/> Grade C: Good		<input type="checkbox"/> Duplicate publication	
<input type="checkbox"/> Grade D: Fair	<input type="checkbox"/> Grade C: A great deal of language polishing	<input type="checkbox"/> Plagiarism	<input type="checkbox"/> Rejection
<input type="checkbox"/> Grade E: Poor		<input checked="" type="checkbox"/> No	<input checked="" type="checkbox"/> Minor revision
	<input type="checkbox"/> Grade D: Rejected	BPG Search:	<input type="checkbox"/> Major revision
		<input type="checkbox"/> The same title	
		<input type="checkbox"/> Duplicate publication	
		<input type="checkbox"/> Plagiarism	
		<input checked="" type="checkbox"/> No	

## COMMENTS TO AUTHORS

I am reviewing this paper as a medical statistician/epidemiologist highly familiar with the conduct of meta-analysis, but with no prior knowledge of factors involved in prognosis for colorectal cancer. I found the paper generally to be a good one, very clearly written in excellent English, and using procedures which mainly seem appropriate. However, I did have some points that require attention. My main issue is with the treatment of heterogeneity. Where estimates are made separately for different levels of a factor, such as study size, study location, gender etc, the main objective should not be to demonstrate whether the hazard ratios (HRs) are statistically significant at each level, but to test whether the hazard ratios vary significantly by level – but such a test is missing. Looking at Table 2, and at the results by number of patients, it is clear from the similarity of the HRs for >100 and <100 that there is no difference by this factor, and emphasising that results are significant for >100 but not for the <100 is off the point. The main message is that there is no evidence of heterogeneity by number of patients, and there is little point in enlarging on that. In fact, there is no evidence of heterogeneity by any of the factors studied, not even for location. The 95% CIs for Asia

are wide and include the overall estimate of 0.69. While you can mention the lack of significance for Asia, this may relate to the high heterogeneity there. Certainly, the estimates by the other factors seem so similar that one need only say that there was no evidence of heterogeneity by these. I also had a problem with the first paragraph of the results section. Thus, it is stated that, of the included articles, 22 enrolled more than 100 patients, and 12 recruited less than 100. However, from Table 1 I make it 24 > 100, 15 < 100 and 1 = 100. Also, it is stated that patients from 7 studies received one of various treatments while in the other 18 studies this was not the case. But 7 + 18 = 25 not 40, as expected. This sentence needs modification. Tests were carried out excluding studies "sequentially". I think "one at a time" is meant, "sequentially" implying that one removed one study, then a second one as well, and so on. Arguably, this is not really required. A comment as to the studies which, when removed, changed the estimate most in each direction would seem adequate. A few other minor points, in the order they appear in the text. 1. Best if possible to avoid abbreviations in the abstract, and only use them in the text, giving the definition when they are first used. Certainly don't, as currently in the abstract, define OS after its use and not define DFS/FRS at all. 2. It is unusual to have the materials and methods last. Logically, they should appear after the introduction (although the journal may have its own specific requirements). 3. In Table 1 "(P/N)" is undefined. 4. At the end of results/methodological quality, there are 19 not 18 studies with a score of 7 or more according to Table 1. 5. At the end of the subgroup analysis section, there is a 0528 missing its decimal point. 6. At the end of the impact section, the final sentence should end "the pooled OR being shown in Table 3". 7. In discussion para 4 line 2, better to say "The relationship between Bcl-2 expression and ...". 8. In the final paragraph of the discussion line 4 "where patients come from ...", not "which patients were come from". 9. The final sentence of the same paragraph should start "Our analysis also found those significant associations only in populations more than 100. This told us ...". Though the whole sentence might disappear for reasons noted above. 10. The final sentence of the materials and methods section could read "Statistical significance was defined as  $p < 0.05$ ".