Reviewer's report

Title: Analysing clinical reasoning characteristics using a combined methods approach.

Version: 5 Date: 17 September 2013

Reviewer: Bernard Charlin

Reviewer's report:

The new version represent a well written paper. It is interesting to read and provides a useful addition in the literature

I only have minor suggestions
in the ABSTRACT:

Background section: authors speak of “the clinical reasoning process”. I wonder if it won’t be appropriate to speak of the clinical reasoning processes. That would underline the complexity of clinical reasoning and the necessity of having a battery of test to measure them.

Result section: The way authors present the values of Cronbach's alpha coefficient is misleading. It suggest that the alpha is .36 for the SCT and 0.61 for the CRP, while in reality the value is 0.60 for SCT2 (the difference of reliability for SCT 1 and 2 is a curious phenomenon that may need some explanation in the discussion section)

METHOD

Who are the experts is not clear. 17 GPs were approached and 12 answered. Do they represent only a level of competence or do they also are those who make the panels for construction of the answer keys for both the CRP and SCT?

The scoring scheme sees to provide points only to the “best answer” of the experts. While described in the literature (Blandt 2005, or Reed 2011) this is an unusual way to score SCTs so this should be specified.

DISCUSSION

Within the discussion it might be useful to specify that the Cronbach value is reach in 30’ of testing time for SCT while it takes 90’ for the CRPs. As mentioned above, the difference of reliability for SCT 1 and 2 is a curious phenomenon that may need some explanation

The scoring scheme used for SCT may explain the weak discriminatory power of the SCT. SCT allows to test in context of uncertainty, a situation in which experienced physicians are favored. This effect may be attenuated if only "best answers" are taken into account.

Finally I don’t understand why the Diagnostic thinking inventory is presented as a kind of gold standard for the quality of clinical reasoning. This is not supported by
data in the literature.