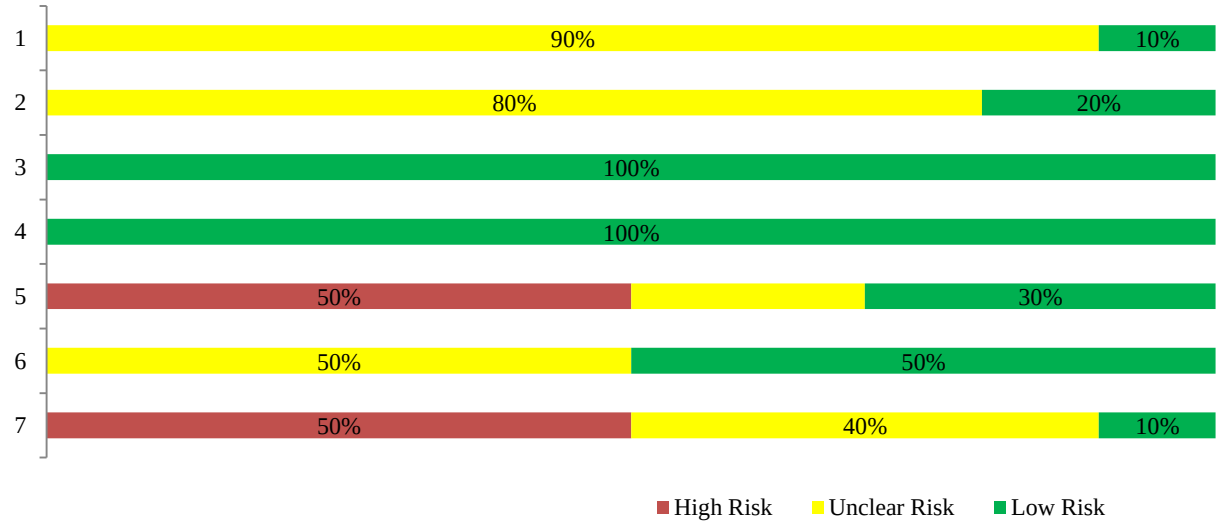
Appendix 1: Aggregate of appraisal of risk of bias of the included studies using the Cochrane risk-of-bias tool [12]



Items:

1. Random sequence generation
2. Allocation concealment
3. Blinding of participants and personnel
4. Blinding of outcome assessment
5. Incomplete outcome data
6. Selective reporting
7. Other bias

Appendix 2. Appraisal of reporting of adverse drug reactions, according to McHarm tool [13]

Study	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Fonseca, 2013 [25]	N	N	N	Y	U	U	Y	N	N	U	Y	Y	Y	Y	Y
Gilman, 2013[a] [26]	N	N	N	N	U	U	N	N	N	N	Y	Y	Y	Y	Y
Moses, 2012 [a] [30]	N	N	N	N	N	N	N	N	N	N	Y	Y	Y	Y	N
Nogueira, 2012 [a] [31]	Y	N	N	N	U	U	N	N	U	Y	Y	Y	Y	Y	Y
Violante, 2012 [32]	Y	N	Y	N	U	U	N	N	N	U	Y	Y	Y	Y	U
Owens, 2011 [33]	Y	Y	N	N	Y	N	N	N	N	U	Y	Y	Y	Y	Y
Hermansen, 2007 [34]	Y	N	N	Y	Y	N	Y	N	U	Y	Y	Y	Y	Y	Y

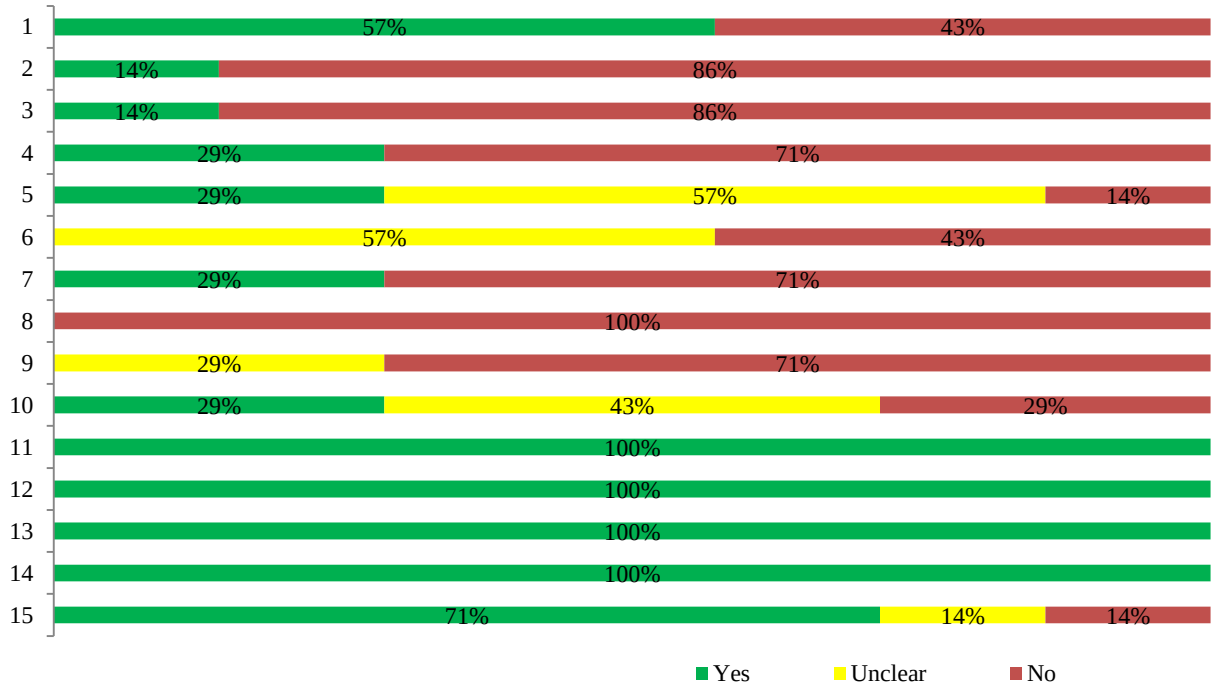
Items:

1. Were harms pre-defined?
2. Were serious events defined?
3. Were severe events defined?
4. Were the numbers of deaths in each study specified?
5. Was the mode of hard collection specified as active?
6. Was the mode of harms collected as passive?
7. Did the study specify who collected the harms?
8. Did the study specify training of background of who ascertained the harms?
9. Did the study specify the timing and frequency of collection of harms?
10. Did the authors use standard scales or checklists for harms?
11. Did the authors specify if the harms reported encompass all the events collected or a selected sample?
12. Was the number of participants that withdrew or were lost to follow up specified for each study group?
13. Was the total number of participants affected by harms specified for each study arm?
14. Did the author(s) specify the number for each type of harmful event for each study group?
15. Did the author(s) specify the type of analyses undertaken for harms data?

N, no; U, unclear; Y, yes.

[a]Unpublished data.

Appendix 3. Aggregate of appraisal of reporting of adverse drug reactions, according to McHarm tool [13]



Items:

1. Were harms pre-defined?
2. Were serious events defined?
3. Were severe events defined?
4. Were the numbers of deaths in each study specified?
5. Was the mode of hard collection specified as active?
6. Was the mode of harms collected as passive?
7. Did the study specify who collected the harms?
8. Did the study specify training of background of who ascertained the harms?
9. Did the study specify the timing and frequency of collection of harms?
10. Did the authors use standard scales or checklists for harms?
11. Did the authors specify if the harms reported encompass all the events collected or a selected sample?
12. Was the number of participants that withdrew or were lost to follow up specified for each study group?
13. Was the total number of participants affected by harms specified for each study arm?
14. Did the author(s) specify the number for each type of harmful event for each study group?
15. Did the author(s) specify the type of analyses undertaken for harms data?