

Reviewer #1:

1. The search methodology should be defined more clearly and should be replicable.

**Response: Done**

2. Abstract should be revised. The background section in the abstract should be reframed. The result section of the abstract is unclear.

**Response: Done**

3. The core-tip section should be more specific and precise.

**Response: Done**

4. The methodology for decision curve analysis can be better described in a tabular/flow-chart/graphical manner.

**Response: Done**

5. The language is sometimes repetitive and needs editing.

**Response: Done**

Reviewer #2:

Major revision: to decrease data heterogeneity by better and more precise inclusion criteria;  
- to involve other studies, that are very important in this area.

**Response: Done, Subgroup study was performed.**

Minor: - too long Introduction; - to describe the implication of presented study to every day endocrine surgeon clinical practice.

**Response: Without a doubt, the introduction is long, however, it contextualizes the real objective of the research and analysis carried out.**

Reviewer #3:

1. According to the results of the meta-analysis, there was significant heterogeneity. Please point out the underlying reasons for the heterogeneity. Sensitivity analysis and subgroup analysis are suggested to be conducted to further address the heterogeneity.

**Response: Done, Subgroup study was performed.**

2. The overall quality of included studies is low, which indicates a high risk of bias.

Minor concerns:

1. The titles of the forest plots are not comprehensive.

**Response: were corrected**

2. The titles of other figures should be revised, such as "Figure 9. Net Benefit interventions avoided, in study patterns."

**Response: Done**

*Science editor:*

Strengths: With a clinical decision-making orientation, using analytical methods of Bayesian analysis, net benefit calculation, etc, different from other studies.

Weaknesses: Some articles with small sample of malignancy, which may cause bias and affect the analysis. Some analytical methods are doubtful, e.g., the determination of overall prevalence of thyroid cancer. The inference (all thyroid nodules which are not completely cystic, regardless of their size, should undergo FNAB) from the results of present study is not consistent with the clinical management of thyroid nodules.

Some specific concerns:

Page 3: The Core tip needs to be rephrased. Page 3: Of "Therefore, we conclude that all thyroid nodules which are not completely cystic, regardless of their size, should undergo FNAB." It's not adequate, and should be deleted or to be rephrased; so does for the Conclusion part at the end of the paper.

**Response: were corrected**

Page 14: Of "The results found in this study show an overall prevalence of cancer of 46.6%." It's derived from that in Page 12: "The data from 8,445 thyroid nodules was obtained, of which 3,937 (46.62%) were malignant and 4,508 (53.38%) were benign. The average size of the tumors was 18.5 mm (5mm to 71 mm)." . It's not right to calculate the overall prevalence of cancer using the above method.

**Response: were corrected**

Page 14: Of "the last line: in particular sensibility and specificity," The word sensibility should be sensitivity, and in other places throughout the paper should be changed too.

**Response: Done**