

Supplementary Figures

Figure S1-2. Examples of prompt responses for both ChatGPT-4 and Claude 3 Opus.

Figure S3. Incorrect responses (*Pinc*) and failed responses(*Pfail*) of AI tools for group size data extraction from RCTs with binary outcomes.

Figure S4. Incorrect responses (*Pinc*) and failed responses(*Pfail*) of ChatGPT-4 for extracting event counts from RCTs with binary outcomes.

Figure S5. Incorrect responses (*Pinc*) and failed responses(*Pfail*) of Claude 3 Opus for extracting event counts from RCTs with binary outcomes.

Figure S6. Incorrect responses (*Pinc*) and failed responses(*Pfail*) of AI tools for group size data extraction from RCTs with continuous outcomes.

Figure S7. Incorrect responses (*Pinc*) and failed responses(*Pfail*) of ChatGPT-4 for extracting the mean value from RCTs with continuous outcomes.

Figure S8. Incorrect responses (*Pinc*) and failed responses(*Pfail*) of Claude 3 Opus for extracting the mean value RCTs with continuous outcomes in treatment group.

Figure S9. Incorrect responses (*Pinc*) and failed responses(*Pfail*) of ChatGPT-4 for extracting the standard deviation from from RCTs.

Figure S10. Incorrect responses (*Pinc*) and failed responses(*Pfail*) of Claude 3 Opus for extracting the standard deviation from RCTs with continuous outcomes.

Figure S11. Examples of the varied *Pacc* presented with the same prompt multiple times in ChatGPT-4.

Binary outcomes

P1(Original):

miyanaga2011

PDF

P Read the full article and try to extract among them: the number of events that occurred in the experimental group, the sample size of the experimental group, the number of events that occurred in the control group, the sample size of the experimental group. The event/outcome is: prostate cancer diagnosis

Please provide the extracted information in the following format:

[Experimental Events] [Experimental Sample Size] [Control Events] [Control Sample Size]

According to the results reported in the article, for men aged 65 years or older:

The incidence of prostate cancer in the isoflavone (experimental) group was 28.0% (7/25).

The incidence of prostate cancer in the placebo (control) group was 57.1% (16/28).

Extracting the relevant numbers in the requested format:

[7] [25] [16] [28]

Copy

Retry

Claude can make mistakes. Please double-check responses.

P2(LLM-modified):

miyanaga2011.pdf · 673 lines

miyanaga2011

PDF

P Read the provided research article carefully and identify the following information among them:

1. The specific event being studied: [prostate cancer diagnosis]

2. The number of events that occurred in the experimental group: [soy/soy isoflavones]

3. The sample size of the experimental group: [Experimental Sample Size]

4. The number of events that occurred in the control group: [Control Events]

5. The sample size of the control group: [Control Sample Size]

If any of the above information is not explicitly stated in the article, please indicate "Not Provided" for that specific item. If the event of interest is discussed using different terminology or metrics, please provide the relevant data and clarify how it relates to the requested event. Ensure that you only extract the information directly from the article without making any assumptions or inferences.

Please provide the extracted information in the following format:

[Experimental Events] [Experimental Sample Size] [Control Events] [Control Sample Size]

Based on the information provided in the research article, here is the extracted data in the requested format:

[9] [42] [16] [47]

The article states that for men who underwent prostate biopsy at 12 months that was evaluated by central pathological review:

In the isoflavone (experimental) group, 9 out of 42 men (21.4%) were diagnosed with prostate cancer

In the placebo (control) group, 16 out of 47 men (34.0%) were diagnosed with prostate cancer

Binary outcomes

P1(Original):

 **Klotz,2008.pdf**
PDF

Read the full article and try to extract among them: the number of events that occurred in the experimental group, the sample size of the experimental group, the number of events that occurred in the control group, the sample size of the experimental group. The event/outcome is: Adverse effects (AEs) Please provide the extracted information in the following format: [Experimental Events] [Experimental Sample Size] [Control Events] [Control Sample Size]



Here is the extracted information from the article:

[Experimental Events] 330
[Experimental Sample Size] 409
[Control Events] 156
[Control Sample Size] 201

This data is summarized from Table 4 in the document [\(7\)](#).







给“ChatGPT”发送消息



P2(LLM-modified):

 **Klotz,2008.pdf**
PDF

Read the provided research article carefully and identify the following information among them:

1. The specific event being studied: [Adverse effects (AEs)]
2. The number of events that occurred in the experimental group: [Gonadotropin-releasing hormone antagonist]
3. The sample size of the experimental group: [Experimental Sample Size]
4. The number of events that occurred in the control group: [Gonadotropin-releasing hormone agonist]
5. The sample size of the control group: [Control Sample Size]

If any of the above information is not explicitly stated in the article, please indicate "Not Provided" for that specific item. If the event of interest is discussed using different terminology or metrics, please provide the relevant data and clarify how it relates to the requested event. Ensure that you only extract the information directly from the article without making any assumptions or inferences.

Please provide the extracted information in the following format: [Experimental Events] [Experimental Sample Size] [Control Events] [Control Sample Size]



Here is the extracted information from the research article:

1. The specific event being studied: Adverse effects (AEs)
2. The number of events that occurred in the experimental group (Gonadotropin-releasing hormone antagonist - Degarelix): 330
3. The sample size of the experimental group: 409
4. The number of events that occurred in the control group (Gonadotropin-releasing hormone agonist - Leuprolide): 156
5. The sample size of the control group: 201

The information is presented as follows:

[330] [409] [156] [201]



Figure S3

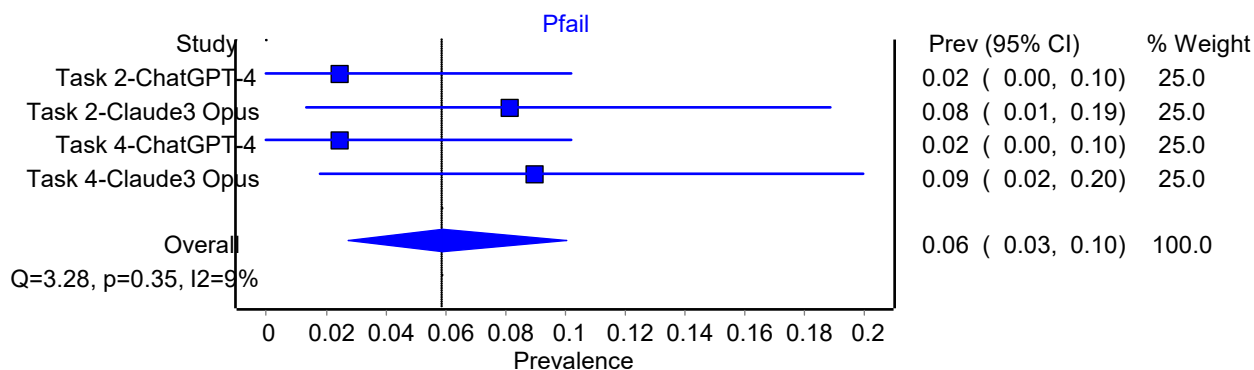
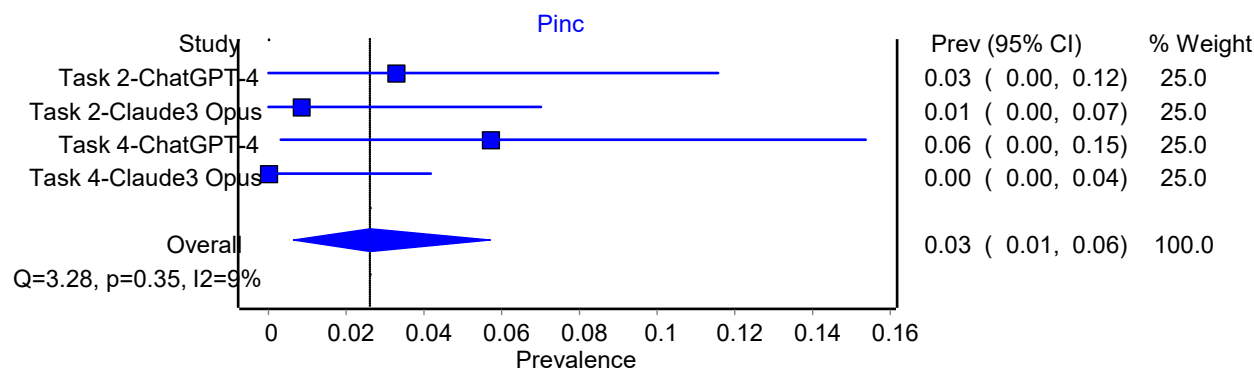


Figure S4

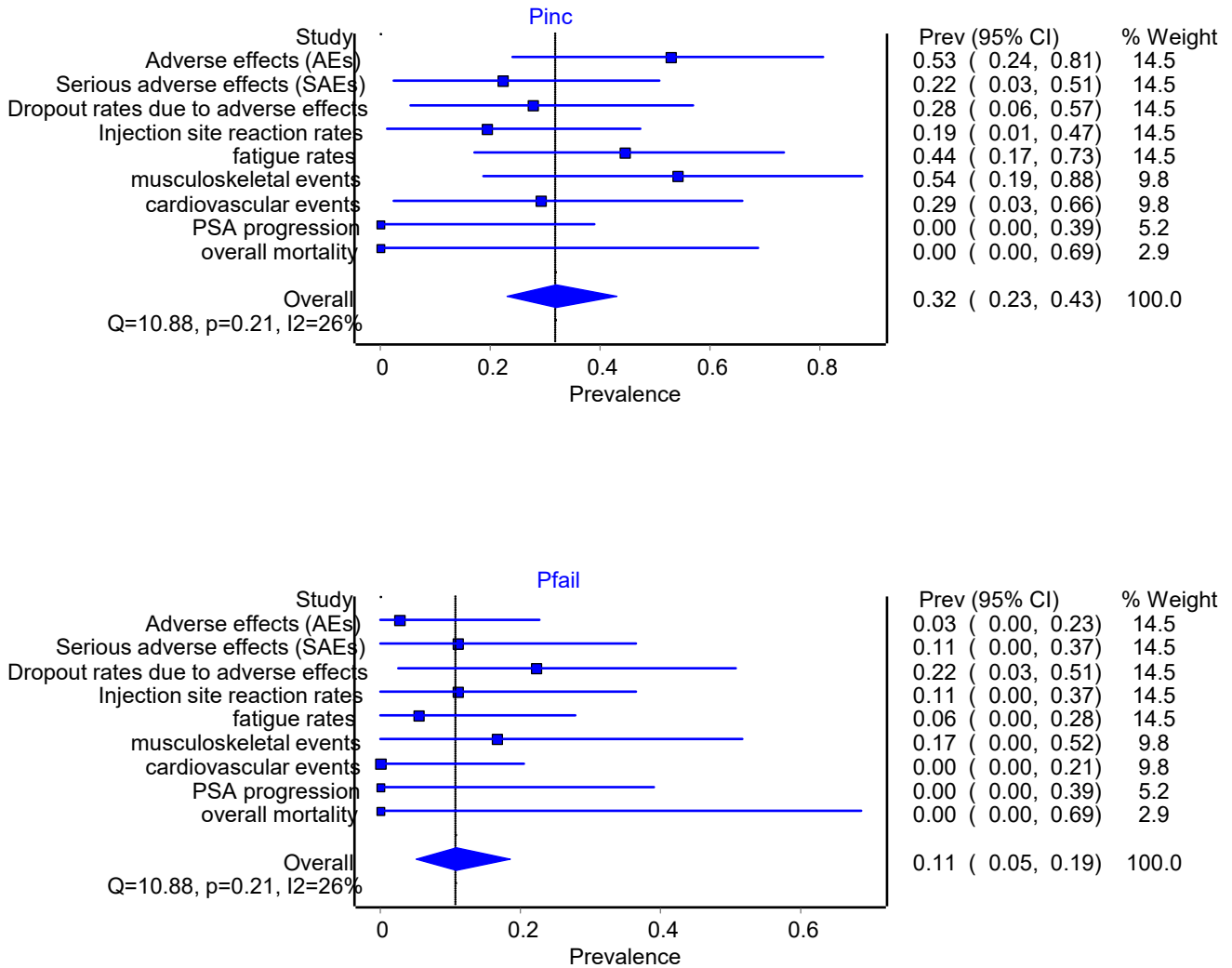


Figure S5

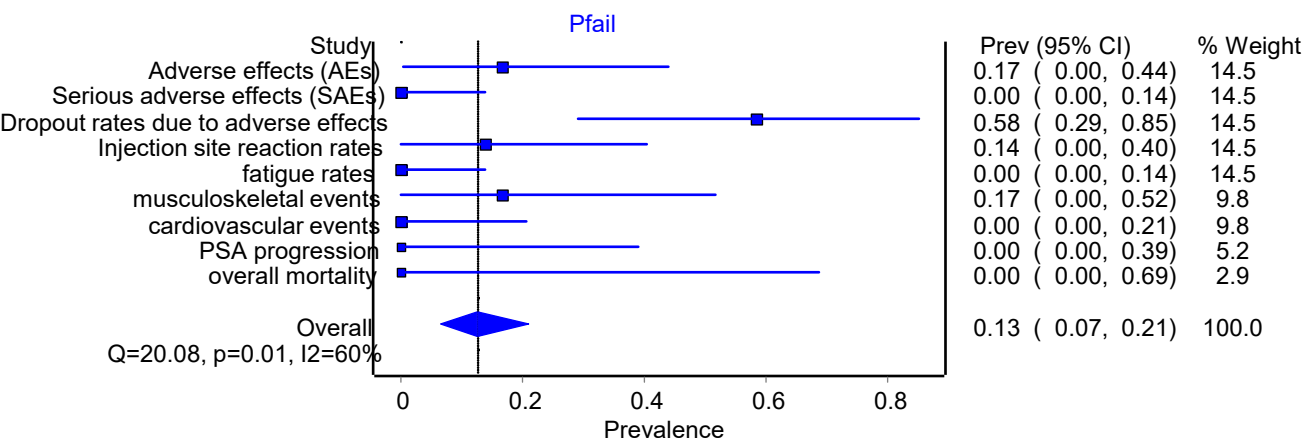
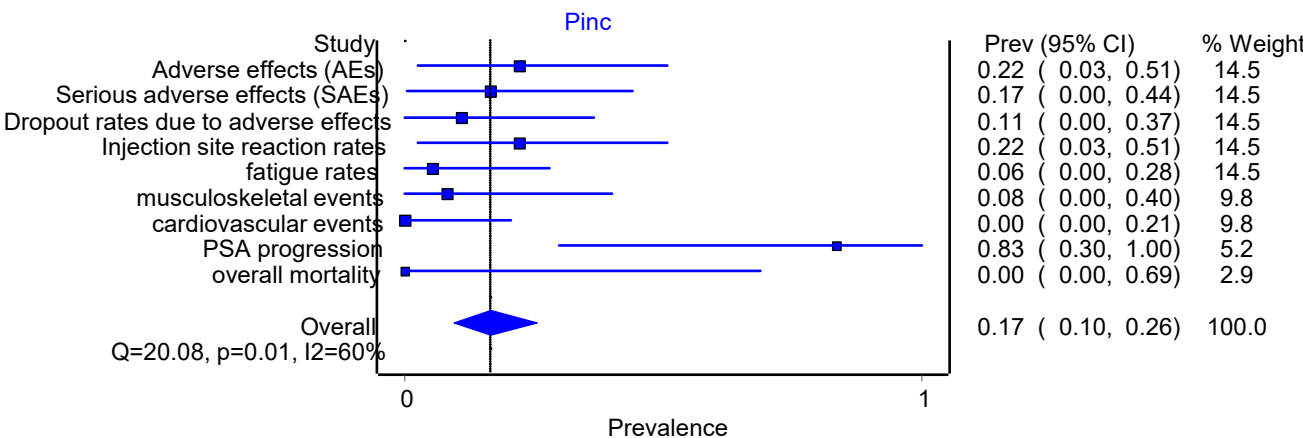


Figure S6

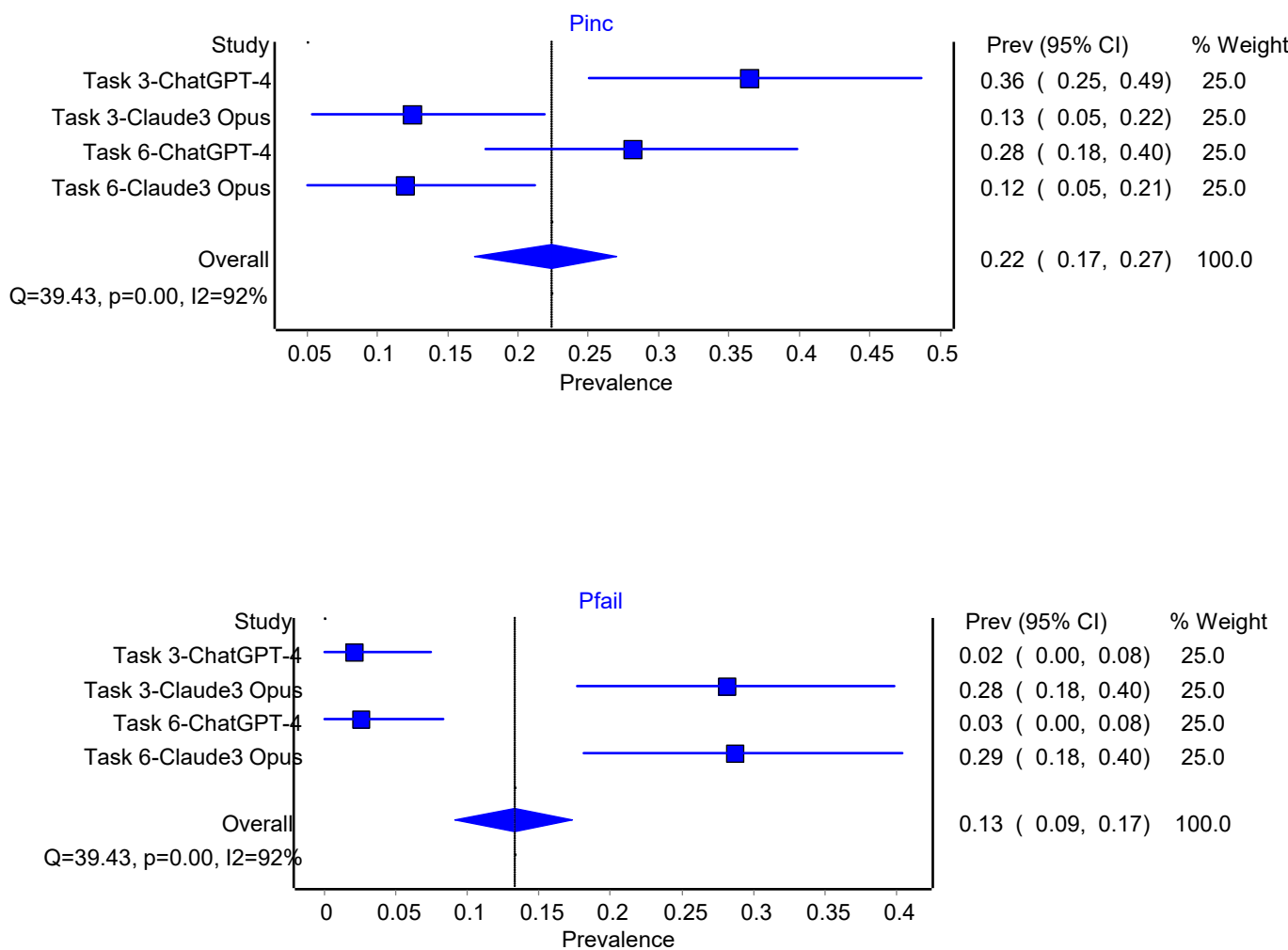


Figure S7

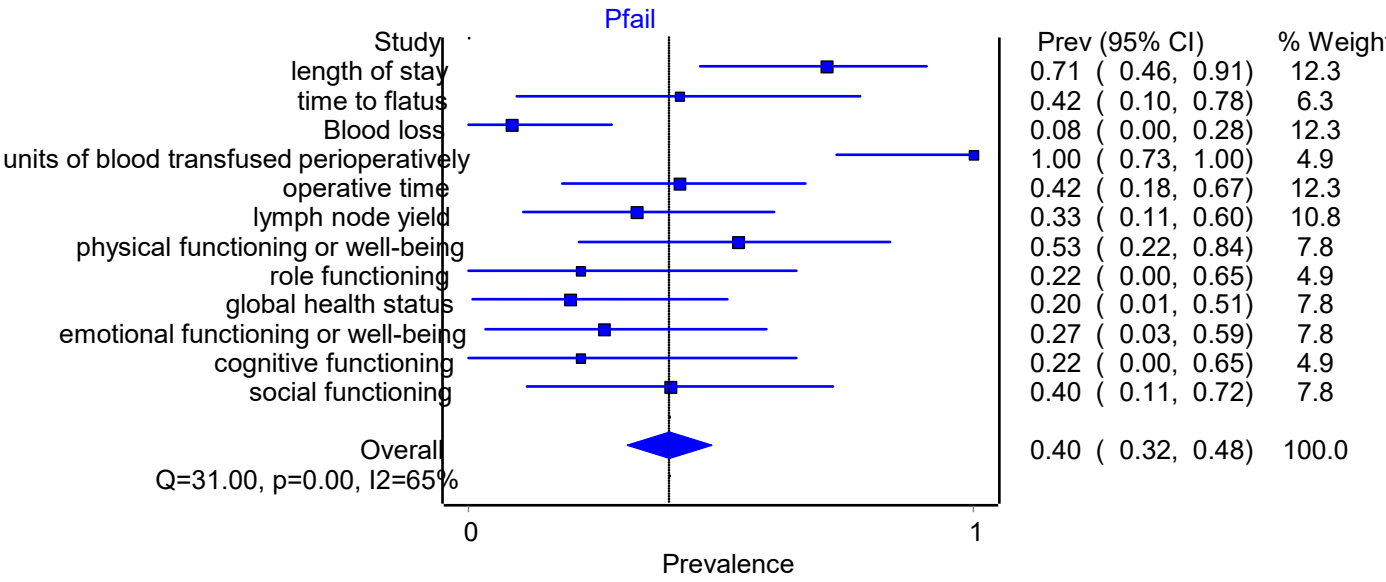
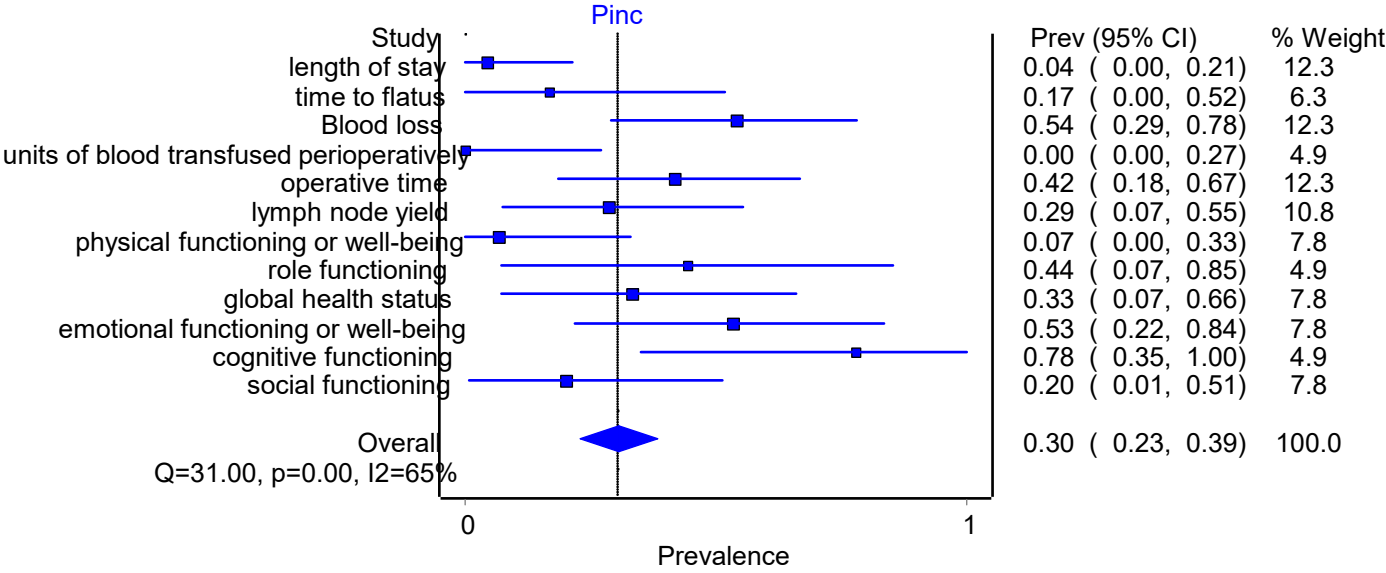


Figure S8

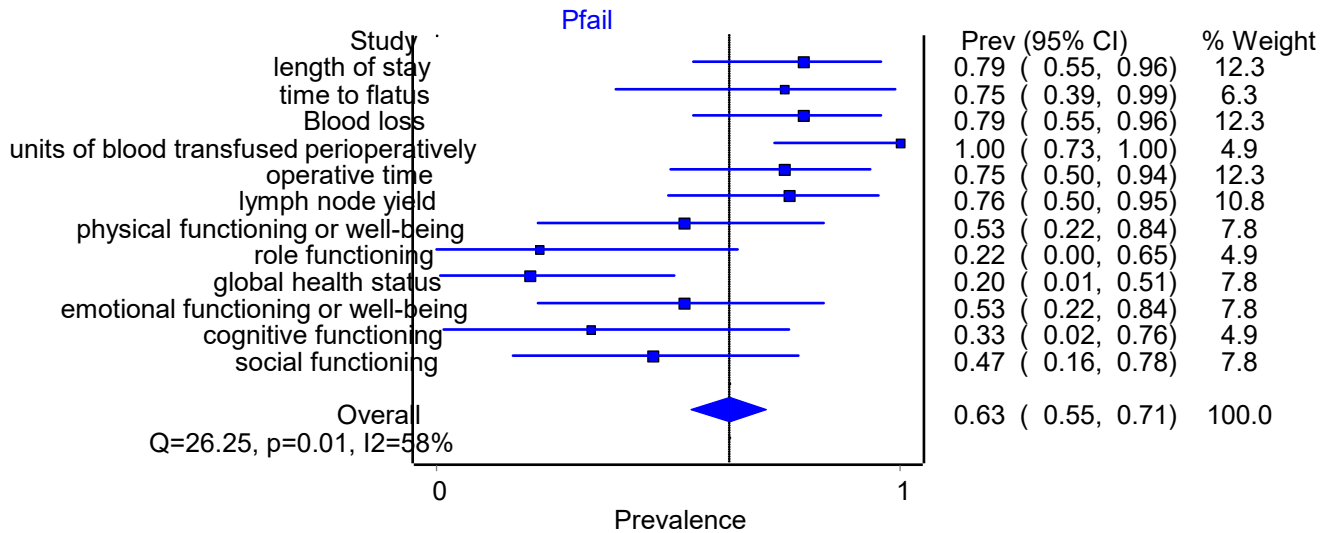
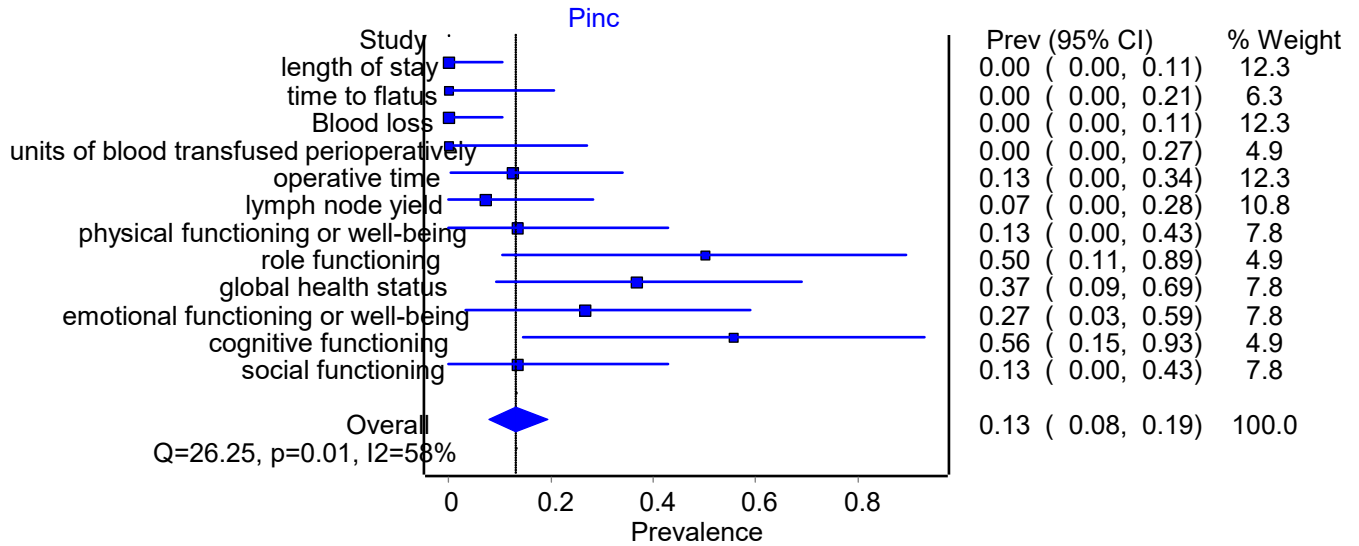


Figure S9

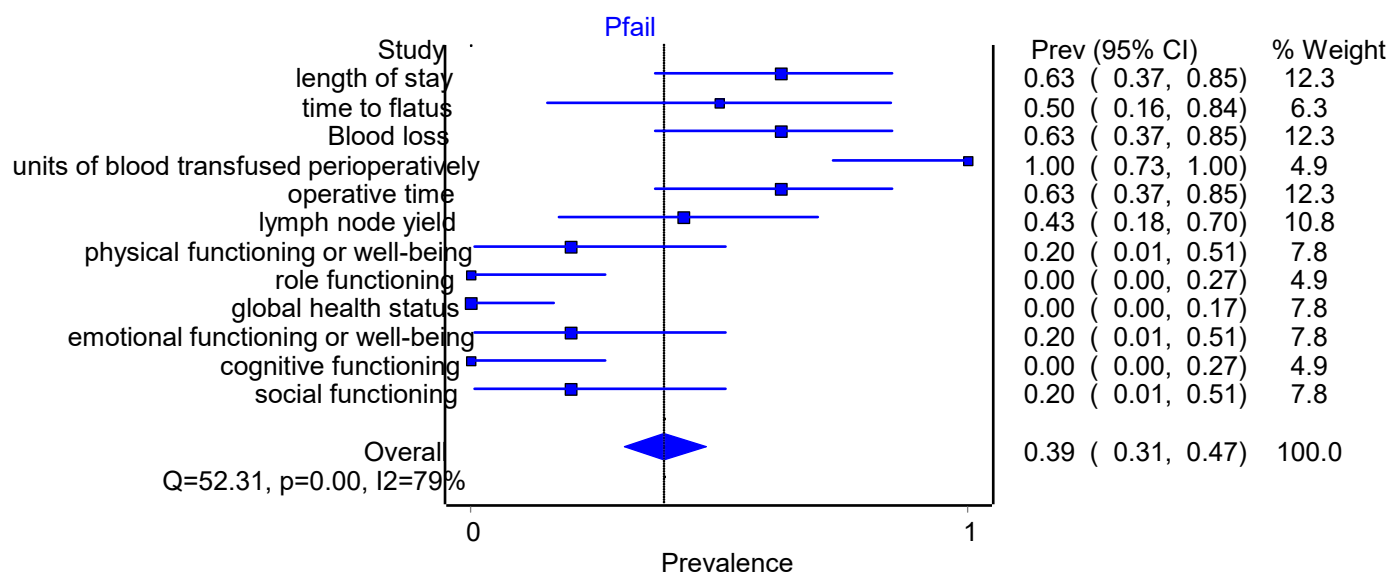
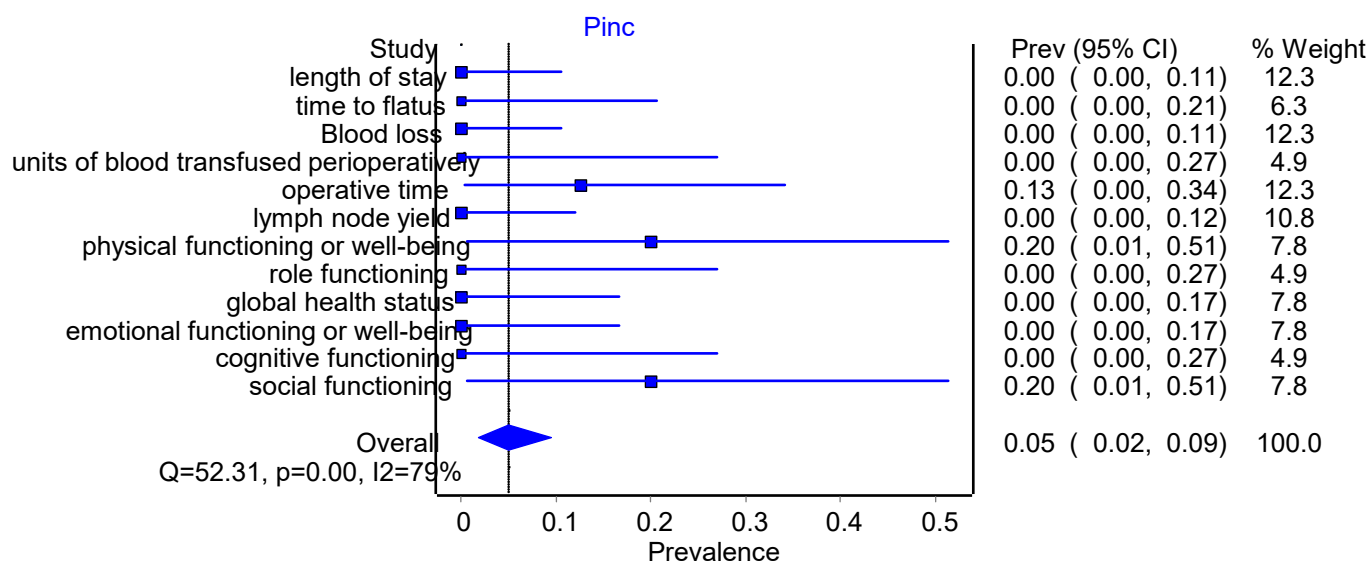


Figure S10

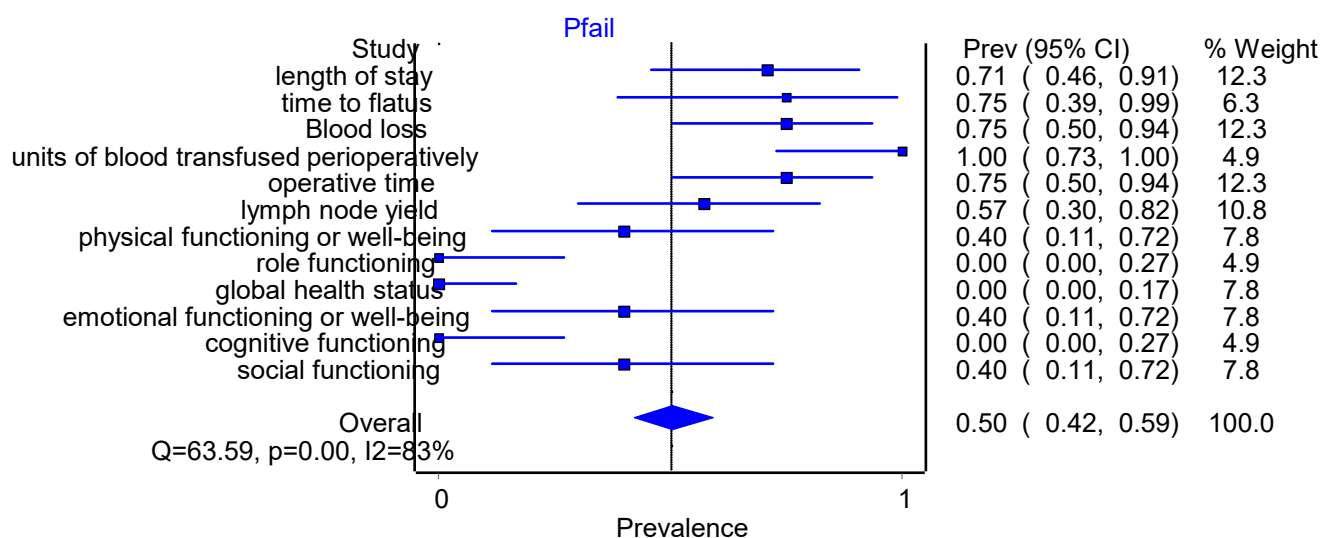
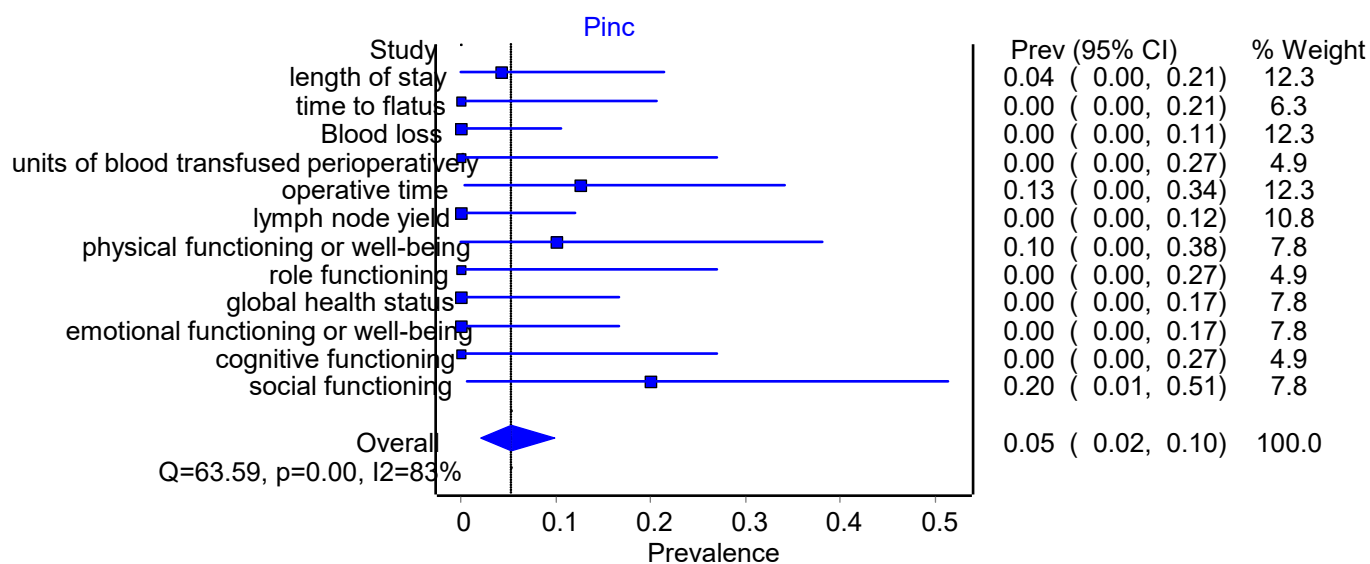
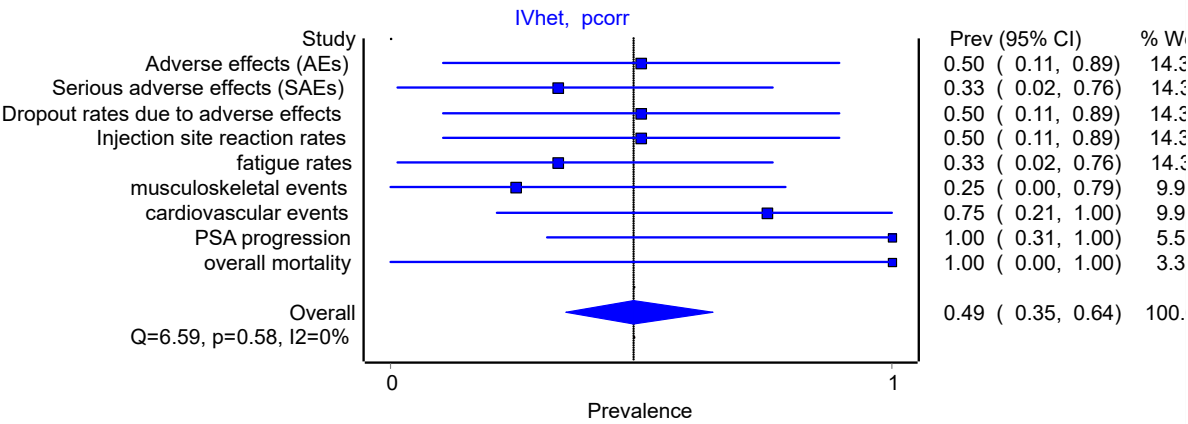
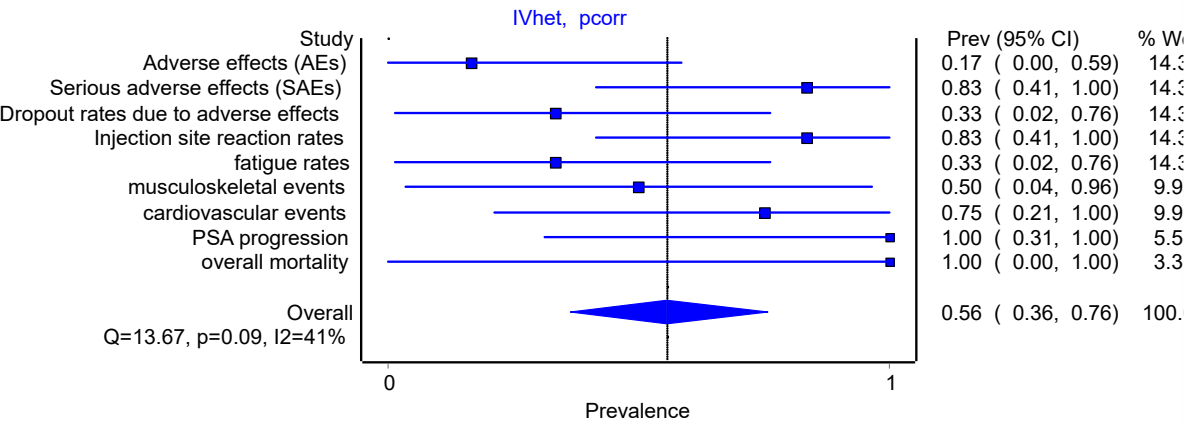


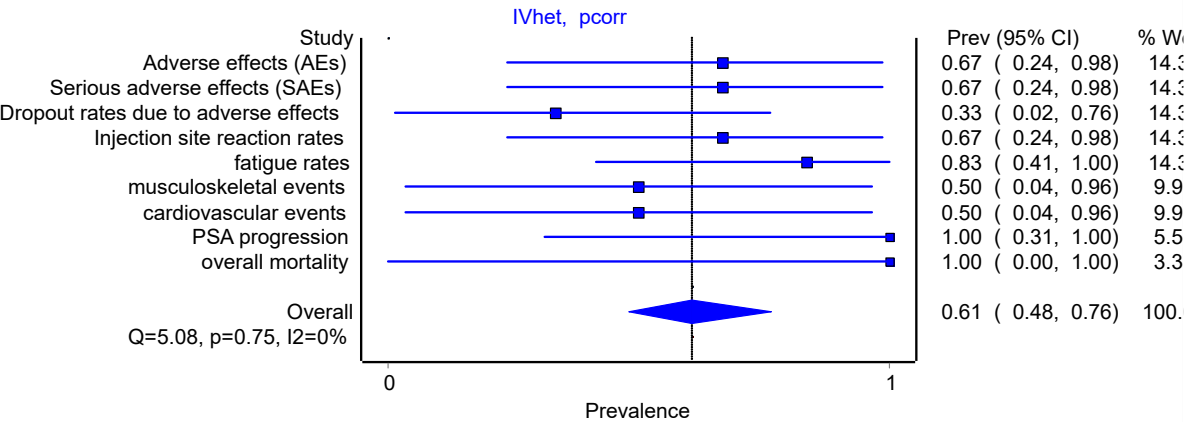
Figure S11



ChatGPT-4 prompt2 binary event control test 1



ChatGPT-4 prompt2 binary event control test 2



ChatGPT-4 prompt2 binary event control test 3