Filled Checklist: DEAL

1. Title and abstract

The title of the article is not explicitly stated in the provided text. The abstract is also not provided, but it can be inferred from the introduction that the study proposes a new model for radiology report retrieval using sentence embeddings and transformer-based language models. The authors claim that their model outperforms existing models on various benchmarks and has potential applications in clinical decision support systems.

2. Introduction: Rationale and objectives

The introduction is not explicitly mentioned in the provided text snippet.

3. Methods: Steering group

Not applicable as there is no mention of a steering group in the provided text.

4. Methods: Facilitators

The methods used to facilitate the study are not explicitly mentioned in the provided text.

5. Methods: Delphi panel selection and recruitment

Not applicable. The provided text does not mention a Delphi panel or its selection/recruitment process.

6. Methods: Initial item generation process

The initial item generation process is not explicitly described in the provided text.

7. Methods: Questionnaire design and pre-testing

There is no mention of questionnaire design or pre-testing in the provided text.

8. Methods: Number of Delphi rounds

There is no mention of a Delphi round in the provided text.

9. Methods: Definition of consensus and criteria

There is no mention in the provided text about a "Definition of consensus and criteria" or how it was used in the study.

10. Methods: Provision of feedback between rounds

There is no mention of "Methods: Provision of feedback between rounds" in the provided text.

11. Methods: Stopping criteria for rounds

There is no mention of stopping criteria for rounds in the provided text.

12. Methods: Statistical analysis

The statistical analysis was performed by authors C.H.S. and J.H.S.

13. Results: Panel recruitment and retention

The provided text does not mention "Panel recruitment and retention" as a topic or provide any relevant information about it. Therefore, there is no answer to this question based on the given context.

14. Results: Round 1 participation and item ratings/responses

Unfortunately, there is no information in the provided text about "Round 1 participation and item ratings/responses". The text appears to be a scientific article discussing a study on developing a model for radiology report retrieval using natural language processing techniques.

15. Results: Round 2 participation and item ratings/responses (and subsequent rounds)

There is no mention of "Round 2 participation and item ratings/responses" in the provided text.

16. Results: Final consensus items/statements

The study demonstrates that a deep learning model can be used to improve radiology report retrieval by adapting transformer-based language models to radiology. The results show that the model outperforms existing keyword search and semantic search methods in solving crossword puzzles using the Google API. Additionally, the study highlights the importance of mitigating bias in radiology machine learning, particularly in data handling and model development. It also emphasizes the need for practical considerations when developing, purchasing, implementing, and monitoring AI tools in radiology. The study's results are based on a large dataset of radiology reports and images, which were used to train and evaluate the deep learning model. The evaluation metrics used include average precision versus area under the ROC curve. Overall, the study provides insights into the potential applications of deep learning in radiology report retrieval and highlights the importance of addressing bias and practical considerations in AI tool development.

17. Discussion: Summary of main findings

The main findings of this study are: * A deep learning model was developed to improve radiology report retrieval by adapting transformer-based language models to radiology. * The model outperformed existing keyword search and semantic search methods in retrieving relevant reports. * The model's performance was evaluated on a large dataset of radiology examinations, with results showing improved recall and precision compared to existing methods. * The study also discussed the importance of mitigating bias in machine learning models for radiology, citing previous studies that highlighted the need for careful consideration of data handling and model development. However, due to the limited context provided, it is unclear what specific findings were made regarding the model's performance on different types of radiology reports or its potential applications in clinical settings.

18. Discussion: Strengths and limitations

The strengths of the model include its ability to adapt transformer-based language models to radiology, improve neural models for radiology report retrieval with lexicon-based automated annotation, and mitigate bias in radiology machine learning. Limitations of the model include: * The potential risk of bias in chest radiography deep learning foundation models. * The need for further study to evaluate the model's performance on diverse patient populations. * The dependence on

large datasets and computational resources. * The possibility of overfitting or underfitting, which could be addressed through more extensive training data or regularization techniques. Additionally, the authors mention that a separate future study will be performed to assess the model's limitations in real-world clinical settings.

19. Discussion: Comparison with existing literature

The discussion mentions comparing the model's performance to other models in radiology and natural language processing. However, it does not explicitly compare the results of this study to existing literature.

20. Discussion: Implications for practice or policy

The implications of this study are: * The proposed model has the potential to improve radiology report retrieval and diagnosis by leveraging semantic search capabilities. * The model can be used as a diagnostic aid for clinicians, potentially reducing errors and improving patient outcomes. * Further research is needed to evaluate the model's performance in real-world clinical settings and to address issues such as bias and generalizability. However, it seems that more context is required to provide a more comprehensive answer.

21. Other Information: Funding sources

The study was supported by the Department of Radiology and Biomedical Imaging, University of California, San Francisco.

22. Other Information: Conflicts of interest

The authors disclose conflicts of interest as follows: - C.H.S.: On the Radiology: Artificial Intelligence Trainee Editorial Board - G.C.: On the Radiology: Artificial Intelligence Trainee Editorial Board - A.D.S.: Associate editor for Radiology - J.H.S.: Early career consultant to the editor for Radiology