Innovative Journal of Pediatrics
Volume 34 · Issue 2 · e142789

OPEN PEER REVIEW

# Innovative Diagnostic Tool: Convolutional Neural Network for Early Fat Malabsorption Detection in Pediatric Patients with Chronic Diarrhea

Emre Kıymık, Erkan Kıymık, Ahmet Basturk

| SUBMITTED | REVISED | ACCEPTED | DOI |
| --- | --- | --- | --- |
| 6 Nov 2023 | 30 Jan 2024 | 1 Mar 2024 | 10.5812/ijp-142789 |

## About this report

This open peer review (OPR) report shows the public comments made on this article during peer review, by the reviewers, the associate editor and the editor-in-chief, for every revision round, together with the authors' replies. Comments appear only with the consent of the people who wrote them. Reviewers who chose to sign their report are named; the others are shown as "Referee". Comments are edited for language only, and confidential remarks to the editors are not included.

More about open peer review: brieflands.com/knowledgebase/opr

## Contents

1 Original submission Associate Editor ..... 2
2 Revision 1 ..... 3
Authors' reply to reviewers ..... 4
Reviewer and editor comments ..... 12

## 1 Original submission

1 comment · 12 Dec 2023

### Associate Editor

MAJOR REVISION 12 Dec 2023

The imbalance in the dataset (19 positive cases out of 162) was addressed using the SMOTE technique. Meanwhile, relying solely on synthetic data to balance the dataset might introduce biases or limitations in the model's performance on real-world data. How have the authors addressed this issue?
This article needs English and grammatical editing.

2

# Revision 1

AUTHORS' RESPONSE TO REVIEWERS

Authors' reply to reviewers

The next pages contain the authors' point-by-point reply to the comments of this revision round, as submitted with the revised manuscript. It is published as the authors wrote it.

*Reviewer 1:*

*This study was conducted with the purpose of INNOVATIVE DIAGNOSTIC TOOL: CONVOLUTIONAL NEURAL NETWORK FOR EARLY FAT MALABSORPTION DETECTION IN PEDIATRIC PATIENTS WITH CHRONIC DIARRHEA is well written, however there are some concerns that should be considered.*

*During the study, it has been repeatedly stated that this study is unique in its kind. These sentences should only be mentioned in the strengths section, and this creates a kind of conflict of interest, especially since the limitations of the study are not mentioned.*

*The references to the uniqueness of the study have been relocated to the Discussion section, aligning with the reviewer's suggestion. Simultaneously, the limitations of the study have been explicitly addressed and incorporated into the discussion, providing a comprehensive and transparent overview of the research*

*Is your study done with the purpose of checking this statistical model or with a clinical purpose? Now, since this is a clinical study, you should discuss the clinical results of the study and compare it with other studies. Please talk about this. For example, what is the status of checking the serum level of micronutrients or liver enzymes in your study? Discuss these.*

*We appreciate the reviewer's insightful comment. Our study serves a dual purpose by validating a unique statistical model and addressing clinically relevant aspects. Due to the innovative nature of our approach, there are currently no directly comparable studies. However, we understand the importance of delving into the clinical implications more comprehensively. In the revised manuscript, we will offer a detailed discussion on the clinical results of our study, emphasizing the significance of our findings. While direct comparisons with other studies may be limited due to our distinctive methodology,*

*What were the criteria and variables you entered into the statistical model based on? Is there a specific reference?*

*The etiology that could cause chronic diarrhea was investigated and blood tests were planned accordingly. Gender, height, weight and age information were included to increase accuracy. This sentences "*The criteria and variables entered into the statistical model were determined based on the investigation of potential etiologies causing chronic diarrhea. To ensure a comprehensive assessment, blood tests were planned accordingly. *" are placed in Dataset Preperation section.*

*For Figure 1, provide more explanations in the caption section.*

*We update as "*Figure 1. Architecture of the Convolutional Neural Network (CNN) Model. The diagram illustrates the key components and layers of the CNN model*"*

*Reviewer 2:*

*Dear authors,*

*I reviewed the manuscript "Innovative Diagnostic Tool: Convolutional Neural Network for Early Fat Malabsorption Detection in Pediatric Patients with Chronic Diarrhea" with great interest. Please see my comments and suggestions as follows:*

*Major comments:*

*- What is the importance and clinical significance of developing this AI-based diagnostic tool?*

*We place the paragraph given below to emphasize advantages and importance of model.*

*"*With the developed deep learning model, fat malabsorption was determined without taking a stool sample. The developed deep learning model offers three important advantages over traditional methods. The first advantage is rapid and early diagnosis. The second advantage is that collecting stool samples can be troublesome and embarrassing for patients, and it offers a much more comfortable analysis considering such situations. The third advantage is cost-effectiveness. The developed CNN model is important step for pediatric gastroenterology, providing an innovative and efficient diagnostic approach to meet the urgent need for accurate and timely diagnosis in children with chronic diarrhea.*"*

*- The study included a relatively small sample size of 100 pediatric patients, impacting the generalizability of the findings. Additionally, the broad inclusion criteria (persistent diarrhea lasting more than four weeks without a prior diagnosis) potentially leads to heterogeneity in patient conditions. How did the authors address this heterogeneity in their primary pool? Also, did the authors calculate the minimum number of required patients before the study? If so, please include the details of the sample size calculation.*

*We appreciate the reviewer's thoughtful consideration of the sample size and heterogeneity concerns. In our study, we acknowledge that the sample size of 100 pediatric patients is relatively small, and we recognize its impact on the generalizability of the findings. To address potential heterogeneity arising from broad inclusion criteria, we employed a systematic approach. The dataset was divided into positive and negative cases of fat malabsorption, with 19 positive and 81 negative cases. Notably, we used the Synthetic Minority Over-sampling Technique (SMOTE) to generate additional positive samples, resulting in a balanced training set of 81 positive and 81*

*negative cases for model development.*

*For the evaluation of model performance, we reserved 30 patients for the test dataset, ensuring that all test data were selected exclusively from the original dataset and not from synthetic data. This approach was adopted to accurately assess the model's performance on real-world scenarios. The test dataset comprised 9 positive and 21 negative cases. Furthermore, we conducted a clinical study involving 27 patients, with 6 testing positive for fat malabsorption. The remaining cases were negative. Notably, our model demonstrated perfect predictions in identifying fat malabsorption in both the test and clinical datasets.*

*While we did not perform a traditional sample size calculation due to the unique aspects of our study, we acknowledge the importance of considering sample size in future investigations. Our study design and robust evaluation methodologies aimed to provide valuable insights into the performance of the model despite the inherent limitations of a relatively small sample size and potential heterogeneity in patient conditions*

*- While Convolutional Neural Networks are powerful for pattern recognition, their 'black-box' nature might limit the interpretability of the model's decisions. Discussing the methods to interpret the outputs, or identifying which features contribute most to the diagnosis, would enhance the study's quality and transparency.*

*To enhance the transparency and interpretability of our model's decisions, we have employed Layer-wise Relevance Propagation (LRP) – a method well-suited for interpreting variants of deep learning. In the Results section, we present the relevance scores of individual features, offering insights into their contributions to the diagnostic process. Moreover, in the Discussion section, we delve into a detailed analysis, identifying and discussing the features that exhibit the most significant impact on the model's diagnostic outcomes. This strategic use of LRP not only contributes to the overall quality of our study but also facilitates a clearer understanding of the key factors influencing the CNN model's decision-making process.*
*We express our sincere gratitude for your valuable insights into the interpretable AI aspect of our study. Your feedback has significantly enhanced the transparency and overall quality of our research.*

*- The imbalance in the dataset (19 positive cases out of 162) was addressed using the SMOTE technique. Meanwhile, relying solely on synthetic data to balance the dataset might introduce biases*

*or limitations in the model's performance on real-world data. How have the authors addressed this issue?*

*We appreciate the reviewer's thoughtful consideration of the dataset imbalance issue. In our study, where we aimed to detect fat malabsorption as positive or negative, we initially encountered a dataset comprising 100 patients with only 19 positive cases. To overcome this imbalance, we strategically employed the Synthetic Minority Over-sampling Technique (SMOTE). Specifically, we generated additional positive samples, resulting in a balanced training set of 81 positive and 81 negative cases.*

*To ensure the robustness of our model and evaluate its real-world performance accurately, we took careful steps in the design of our test and clinical datasets. The test dataset, consisting of 30 patients, was exclusively composed of original data, with 9 positive and 21 negative cases. This deliberate choice aimed to evaluate the model's performance on real-world data, mitigating potential biases introduced by relying solely on synthetic data. Importantly, the model exhibited precise predictions in this evaluation, correctly identifying fat malabsorption in both positive and negative cases.*

*Furthermore, we extended our evaluation to a clinical study involving 27 patients, of which 6 were positive for fat malabsorption. The remaining patients were negative. Notably, our model demonstrated perfect predictions in this clinical dataset as well. These meticulous design choices were implemented to ensure the model's robustness and effectiveness in real-world scenarios, addressing concerns related to biases or limitations introduced by synthetic data during training.*

*- The model validation is clearly described. However, the metrics used for model evaluation and their clinical relevance should be further elaborated. Discussing the sensitivity, specificity, positive predictive value, etc., would provide a clear picture of the real-world utility of the proposed model and its advantages.*

*In response to the reviewer's valuable feedback, we appreciate the suggestion to further elaborate on the metrics used for model evaluation and their clinical relevance. In the Results section, we have presented key metrics such as F1 score, Precision, and Recall, providing insights into the model's performance. To offer a more comprehensive understanding of the real-world utility of our proposed model and its advantages, we have strategically placed additional metrics including sensitivity, specificity, positive predictive value, true positives (tp), and false negatives (fn) in the Discussion section. This intentional placement aims to enhance clarity and facilitate a detailed discussion on the clinical relevance and practical implications of our model's performance.*

*- External validation on an independent dataset and in a clinical setting outside the research environment is highly recommended.*

*Following your advice, we conducted an additional research effort to obtain an external dataset comprising 27 patients, where 6 tested positive for fat malabsorption, and the remaining cases were negative. We have diligently included the results of this external validation in the Results section, providing a transparent account of the model's performance on this independent dataset.*
*Furthermore, to ensure a thorough discussion of the clinical implications and robustness of our findings, we have included detailed insights from this external validation in the Discussion section.*

*- Did the authors identify/address the potential confounding factors in diagnosing fat malabsorption? Statistical adjustments or considerations for confounders should be addressed to ensure robustness in the results.*

*we did not conduct explicit checks for confounding factors. The nature of deep learning involves updating weights, and features deemed confounding or unnecessary would naturally have their weights mitigated during the learning process. However, thanks to your insightful advice, we incorporated the Layer-wise Relevance Propagation (LRP) method in our study. LRP provided relevance scores, shedding light on the effectiveness of various features and factors. This strategic addition allows us to identify which features play the most significant and least significant roles in the diagnostic process, contributing to a more nuanced understanding of potential confounding factors and enhancing the robustness of our results.*

*The discussion section should include more information about the already-developed models, their advantages and disadvantages compared to the current model, and, more importantly, the clinical importance and significance of developing an AI-based diagnostic tool for this purpose.*

*Our AI model for detecting fat malabsorption represents, to the best of our knowledge, the first developed model for this specific purpose. We miswrote this “As far as our knowledge extends, a CNN model or an AI model has been developed for detecting fat malabsorption, utilizing a combination of diverse blood tests and features”, The correct sentence should be “As far as our knowledge extends, a CNN model or an AI model has **not** been developed for detecting fat malabsorption, utilizing a combination of diverse blood tests and features.”.*
*Moreover, we have enriched the discussion section by providing a more detailed exploration of the advantages our model offers over traditional methods used in fat malabsorption detection. This includes considerations of the clinical importance and significance of deploying an AI-based*

*diagnostic tool in pediatric gastroenterology, emphasizing the rapidity, patient comfort, and cost-effectiveness achieved by our innovative CNN model.*

*- The manuscript needs extensive language editing. Also, several tortured phrases, such as “sugar chromatography with hemoglobin” and “patient spots” were spotted, questioning the quality of the study presentation.*

*The phrases and sentences are corrected.*

*Minor comments:*

*- The manuscript should follow the journal’s instructions for authors. Please revise the manuscript’s headings, following the authors’ instructions (https://brieflands.com/journals/iranian-journal-of-pediatrics/knowledgebase/category/tree#research_articles).*

*"When I checked the template, I could not identify any issues with the headings as per the instructions provided in the journal's template (https://brieflands.com/journals/iranian-journal-of-pediatrics/knowledgebase/category/tree#research_articles). Please let me know if there are specific details or sections that need further adjustments for compliance with the journal's instructions."*

*- The abbreviations should be fully defined at the first use, even in the abstract.*

*- Background, pages 5-6: Several paragraphs of the Background section describe the study’s methods. The same goes for the Objectives section in the main text and the Methods section of the Abstract. Please revise these sections accordingly.*

*Objective is updated. Abstract is updated.*

*- Please move all reported results from the “Materials”, and the “Methods” sections to the Results section.*

*We removed these tables to Results section.*

*- The inclusion criteria are well-described. However, consider including more details regarding the exclusion criteria, if any. This would ensure clarity about the sample selection process.*

*- On page 5, the first paragraph states: “In our study, we seek to harness the potential of this cutting-edge technology for the early detection of fat malabsorption in children afflicted with chronic diarrhea.” Please move this sentence to the end of the introduction/objectives.*

*Sentence is moved.*

*- Discussion section, “Comparison with Previous Studies” subsection, first paragraph states, “As far as our knowledge extends, a CNN model or an AI model has been developed for detecting fat malabsorption, utilizing a combination of diverse blood tests and features”. Please cite the proper references and indicate the advantages of the current model.*

*Unfortunately, we miswrote this sentence this should be “As far as our knowledge extends, a CNN model or an AI model has **not** been developed for detecting fat malabsorption, utilizing a combination of diverse blood tests and features”*

*- The potential drawbacks of oversampling techniques (such as SMOTE) and how they might impact model generalization should be discussed among the study limitations.*

*We placed the drawbacks of the SMOTE algorithm in the limitation part of the discussion section, and also we placed how we try to solve the drawbacks of the SMOTE algorithms.*

*- The study lacks a discussion on the practicality and cost-effectiveness of implementing this diagnostic tool in real-world clinical settings.*

**Associate Editor's Comments:**

*Associate Editor 1:*

*The imbalance in the dataset (19 positive cases out of 162) was addressed using the SMOTE technique. Meanwhile, relying solely on synthetic data to balance the dataset might introduce biases or limitations in the model’s performance on real-world data. How have the authors addressed this issue?*

*In our study, where we aimed to detect fat malabsorption as positive or negative, we initially encountered a dataset comprising 100 patients with only 19 positive cases. To overcome this imbalance, we strategically employed the Synthetic Minority Over-sampling Technique (SMOTE). Specifically, we generated additional positive samples, resulting in a balanced training set of 81 positive and 81 negative cases.*

*To ensure the robustness of our model and evaluate its real-world performance accurately, we took careful steps in the design of our test and clinical datasets. The test dataset, consisting of 30 patients, was exclusively composed of original data, with 9 positive and 21 negative cases. This deliberate choice aimed to evaluate the model's performance on real-world data, mitigating potential biases introduced by relying solely on synthetic data. Importantly, the model exhibited*

*precise predictions in this evaluation, correctly identifying fat malabsorption in both positive and negative cases.*

*Furthermore, we extended our evaluation to a clinical study involving 27 patients, of which 6 were positive for fat malabsorption. The remaining patients were negative. Notably, our model demonstrated perfect predictions in this clinical dataset as well. These meticulous design choices were implemented to ensure the model's robustness and effectiveness in real-world scenarios, addressing concerns related to biases or limitations introduced by synthetic data during training*

*This article needs English and grammatical editing.*

Grammar check of the manuscript was done.

## 2 Revision 1

0 comments

*No public comments in this round.*