

1. Introduction

This document tests OCR fallback functionality. It contains a mix of pages with native text extraction and pages

 Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolor

2. Methods

This page contains the methods section. It has been converted to an image to simulate a scanned document.

The experimental procedure involved multiple steps including data collection, preprocessing, and statistical analysis.

3. Results

The results of our analysis are presented below. This page uses native text extraction and does not require

Table 1 shows the primary outcomes of the study. The intervention group showed significant improvement.

This page contains a figure that has been scanned.

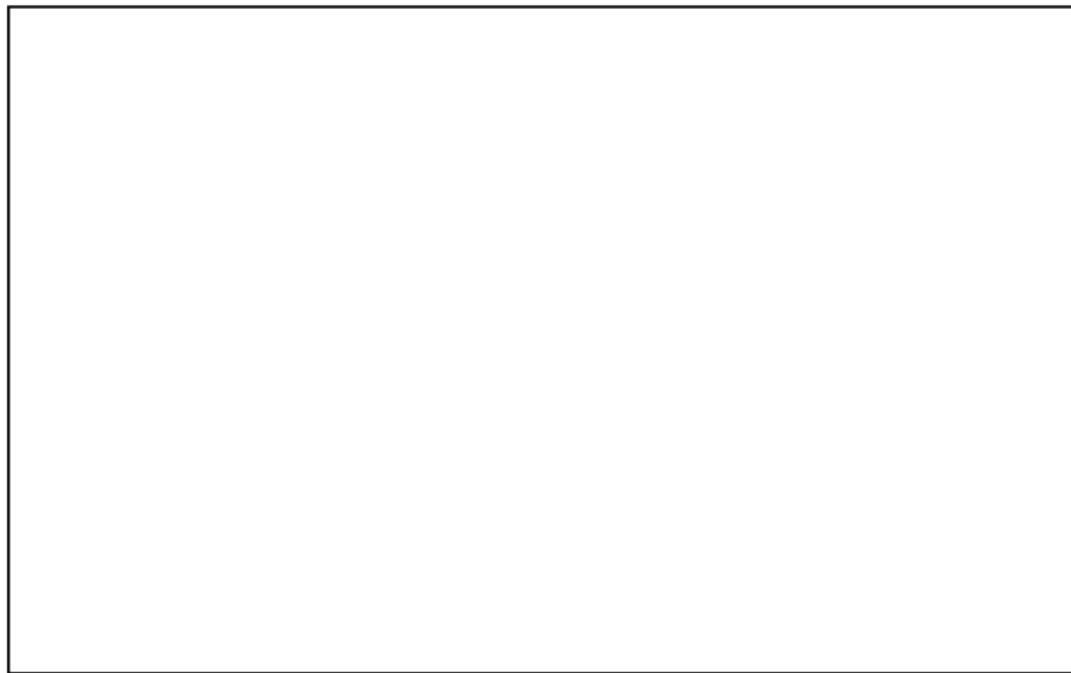


Figure 1: Experimental results showing treatment effect

4. Discussion

Our findings demonstrate the effectiveness of the approach. This page contains native text that should be excluded from the final document.

Future work should investigate the generalizability of these findings to broader populations and clinical settings.

Left Column

This is the left column
of a two-column layout.
Text continues here with
multiple lines of content.
The extraction should
handle this layout.

Right Column

This is the right column
of the same layout.
Reading order may vary
depending on extraction.
Both columns should be
extracted properly.