A Manuscript Microscope 'Sentence Audit' of:

Multi-Modality Machine Learning Predicting Parkinson's Disease

What is the Manuscript Microscope Sentence Audit?

The Manuscript Microscope Sentence Audit is a research paper introspection system that parses the text of your manuscript into minimal sentence components for faster, more accurate, enhanced proofreading.

Why use a Sentence Audit to proofread your manuscript?

- Accelerated Proofreading: Examine long technical texts in a fraction of the usual time.
- Superior Proofreading: Detect subtle errors that are invisible to traditional methods.
- Focused Proofreading: Inspect each individual sentence component in isolation.
- Reliable Proofreading: Ensure every single word of your manuscript is correct.
- Easier Proofreading: Take the hardship out of crafting academic papers.

Bonus 1: Improved Productivity: Rapidly refine rough drafts to polished papers.
Bonus 2: Improved Authorship: Cultivate a clear, concise, consistent, writing style.
Bonus 3: Improved Reputation: Become known for rigorously precise publications.

Manuscript Source: https://www.biorxiv.org/content/10.1101/2021.03.05.434104v1

Manuscript Authors: Mary B. Makarious, Hampton L. Leonard, Dan Vitale, Hirotaka Iwaki, Lana Sargent, Anant Dadu, Ivo Violich, Elizabeth Hutchins, David Saffo, Sara Bandres-Ciga, Jonggeol Jeff Kim, Yeajin Song, Matt Bookman, Willy Nojopranoto, Roy H. Campbell, Sayed Hadi Hashemi, Juan A. Botia, John F. Carter, Melina Maleknia, David W. Craig, Kendall Van Keuren-Jensen, Huw R. Morris, John A. Hardy, Cornelis Blauwendraat, Andrew B. Singleton, Faraz Faghri & Mike A. Nalls

Audit Date: 25/03/21 **Audit Identifier:** 023Y9H82SOY80QB **Code Version:** 3.6

Features of the Sentence Audit:

The Sentence Audit combines two complementary proofreading approaches:

1. Each sentence of your text is parsed and displayed in isolation for focused inspection.
2. Each individual sentence is further parsed into Minimal Sentence Components for a deeper review of the clarity, composition and consistency of the language you used.

The Minimal Sentence Components shown are the smallest coherent elements of each sentence of your text as derived from it's conjunctions, prepositions and selected punctuation symbols (i.e. commas, semicolons, round and square brackets).

The combined approaches ensure easier, faster, more effective proofreading.

Comments and Caveats:

• The sentence parsing is achieved using a prototype natural language processing pipeline written in Python and may include occasional errors in sentence segmentation.

• Depending on the source of the input text, the Sentence Audit may contain occasional html artefacts that are parsed as sentences (E.g. "Download figure. Open in new tab").

• Always consult the original research paper as the true reference source for the text.

Contact Information:

To get a Manuscript Microscope Sentence Audit of any other research paper, simply forward any copy of the text to John.James@OxfordResearchServices.com.

All queries, feedback or suggestions are also very welcome.

Research Paper Sections:

The sections of the research paper input text parsed in this audit.

Section No.	Headings	Sentences
Section: 1	SUMMARY	28
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0
N/A	--- --- --- --- --- --- ---	0

The Sentence Audit Of The Research Paper

Title **Multi-Modality Machine Learning Predicting Parkinson's Disease**

S1 [001] **SUMMARY**

S1 [002] Background

Background

S1 [003] Personalized medicine promises individualized disease prediction and treatment.

Personalized medicine promises individualized disease prediction ...

... and treatment.

S1 [004] The convergence of machine learning (ML) and available multi-modal data is key moving forward.

The convergence ...

... of machine learning ...

... (ML) ...

... and available multi-modal data is key moving forward.

S1 [005] We build upon previous work to deliver multi-modal predictions of Parkinson's Disease (PD).

We build ...

... upon previous work ...

... to deliver multi-modal predictions ...

... of Parkinson's Disease ...

... (PD).

S1 [006] Methods

Methods

S1 [007] We performed automated ML on multi-modal data from the Parkinson's Progression Marker Initiative (PPMI).

We performed automated ML ...

... on multi-modal data ...

... from the Parkinson's Progression Marker Initiative ...

... (PPMI).

S1 [008] After selecting the best performing algorithm, all PPMI data was used to tune the selected model.

After selecting the best performing algorithm, ...

... all PPMI data was used ...

... to tune the selected model.

S1 [009] The model was validated in the Parkinson’s Disease Biomarker Program (PDBP) dataset.

The model was validated ...
... in the Parkinson’s Disease Biomarker Program ...
... (PDBP) ...
... dataset.

S1 [010] Finally, networks were built to identify gene communities specific to PD.

Finally, ...
... networks were built ...
... to identify gene communities specific ...
... to PD.

S1 [011] Findings

Findings

S1 [012] Our initial model showed an area under the curve (AUC) of 89.72% for the diagnosis of PD.

Our initial model showed an area ...
... under the curve ...
... (AUC) ...
... of 89.72% ...
... for the diagnosis ...
... of PD.

S1 [013] The tuned model was then tested for validation on external data (PDBP, AUC 85.03%).

The tuned model was then tested ...
... for validation ...
... on external data ...
... (PDBP, ...
... AUC 85.03%).

S1 [014] Optimizing thresholds for classification, increased the diagnosis prediction accuracy (balanced accuracy) and other metrics.

Optimizing thresholds ...
... for classification, ...
... increased the diagnosis prediction accuracy ...
... (balanced accuracy) ...
... and other metrics.

S1 [015] Combining data modalities outperforms the single biomarker paradigm.

Combining data modalities outperforms the single biomarker paradigm.

S1 [016] UPSIT was the largest contributing predictor for the classification of PD.

UPSIT was the largest contributing predictor ...
... for the classification ...
... of PD.

S1 [017] The transcriptomic data was used to construct a network of disease-relevant transcripts.

The transcriptomic data was used ...

... to construct a network ...

... of disease-relevant transcripts.

S1 [018] Interpretation

Interpretation

S1 [019] We have built a model using an automated ML pipeline to make improved multi-omic predictions of PD.

We have built a model ...

... using an automated ML pipeline ...

... to make improved multi-omic predictions ...

... of PD.

S1 [020] The model developed improves disease risk prediction, a critical step for better assessment of PD risk.

The model developed improves disease risk prediction, ...

... a critical step ...

... for better assessment ...

... of PD risk.

S1 [021] We constructed gene expression networks for the next generation of genomics-derived interventions.

We constructed gene expression networks ...

... for the next generation ...

... of genomics-derived interventions.

S1 [022] Our automated ML approach allows complex predictive models to be reproducible and accessible to the community.

Our automated ML approach allows complex predictive models ...

... to be reproducible ...

... and accessible ...

... to the community.

S1 [023] Funding

Funding

S1 [024] National Institute on Aging, National Institute of Neurological Disorders and Stroke, the Michael J.

National Institute ...

... on Aging, ...

... National Institute ...

... of Neurological Disorders ...

... and Stroke, ...

... the Michael J.

S1 [025] Fox Foundation, and the Global Parkinson’s Genetics Program.

Fox Foundation, ...

... and the Global Parkinson’s Genetics Program.

S1 [026] Evidence before this study

Evidence ...

... before this study

S1 [027] Prior research into predictors of Parkinson’s disease (PD) has either used basic statistical methods to make predictions across data modalities, or they have focused on a single data type or biomarker model.

Prior research ...

... into predictors ...

... of Parkinson’s disease ...

... (PD) ...

... has either used basic statistical methods ...

... to make predictions ...

... across data modalities, ...

... or they have focused ...

... on a single data type ...

... or biomarker model.

S1 [028] We have done this using an open-source automated machine learning (ML) framework on extensive multi-modal data, which we believe yields robust and reproducible results.

We have done this ...

... using an open-source automated machine learning ...

... (ML) ...

... framework ...

... on extensive multi-modal data, ...

... which we believe yields robust ...

... and reproducible results.

End of Sample Audit

This is a truncated Manuscript Microscope Sample Audit.

To get the full audit of this text (or any other research paper),
forward a copy of the research paper to John James at
John.James@OxfordResearchServices.com
