The COVID-19 pandemic’s deterioration of psychological wellbeing in certain populations is well documented.

This paper aims to fill gaps in the literature by reviewing and comparing mental health literature from lower middle-income countries (LMIC) and non-LMICs using natural language processing (NLP).

The created data visualizations demonstrate that the mental health literature remains similar across demographics and study areas with small but notable differences in demographics and reported impact/attitude on mental health and the pandemic in LMICs: a more “negative” response to the pandemic and a focus on younger demographics.

Due to the ongoing nature of the pandemic and the lack of published literature from LMICs, a complete picture of the overall impact on psychological wellbeing cannot be captured. Thus, future research remains to be done.

Research Goal
1. To characterize and compare the existing literature on mental health and psychological wellbeing across different demographics and study areas
   1. How do populations in LMICs compare to populations in more developed nations as reflected by the published mental health literature?

Future Plans
  • Creation of an accurate sentiment analysis classifier capable of labelling data automatically
  • Expanding the dataset
  • Tracking sentiment across time as vaccination campaigns roll out across the world

Acknowledgement
  • The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any agency of the U.S. government. Assumptions made within the analysis are not a reflection of the position of any U.S. government entity. This work is supported by the U.S. Department of State and NSF award numbers 1841403 and 1841520.

Results & Analysis
• The nodes “high”, “anxiety”, “stress”, and “depression” form a tight cluster, indicating a close relationship with one another across all the data.
• The “resilience” node forms a close network with “social”, “emotional”, and “support”, indicating a correlation between the nodes.
• The LMIC data indicates a higher prevalence of terms such as “fear”, “threat”, “tension”, and “gender” whereas those terms don’t occur commonly in the Non-LMIC data.
• The results of the tf-idf analysis indicate that the terms “family”, “gender”, and “differential” are more relevant in the LMIC data and in fact do not appear in the most common frequencies in the Non-LMIC Data.

Conclusions
• Researchers and policymakers must reconsider outreach to target demographics during the pandemic in LMICs
• More representative data of LMICs would allow for better characterization
• Current measures must be revaluated and assessed for possible negative impacts against vulnerable groups in LMIC populations