Associate Editor
Comments to the Author:
The paper has been reviewed by two experts in neuroscience and one expert in causal inference. The reviews are mixed. Upon my examination, I found the manuscript difficult to understand and identified several substantial shortcomings.

1. Initially, the introduction filled me with anticipation for the stochastic intervention, as I expected a significant contribution on utilizing randomized experiments to confirm genuine causal relationships. However, my enthusiasm diminished after I finished reading the paper, as it relies on observational studies and two hypothetical interventions, lacking empirical validation. Consequently, the practical applicability of the mediation framework and the proposed interventions remains questionable, especially since the results in real-world data fail to provide evidence of the framework and hypothesis.

2. The paper focuses too much on methodologies and theories, which diverge from the core interests of the JASA ACS. Moreover, the presented methods and theories are rather conventional in the domain of causal inference, offering limited novelty. I would suggest moving all the theories and part of the method to the supplementary material. Despite the authors present their work as practical in application, it essentially serves as a discourse on causal methodology, using imaging applications as a demonstration of the proposed method. The absence of empirical data analysis to validate the causal mediation framework and essential causal assumptions (A1, A2, A3) undermines its eligibility for publication in JASA ACS.

3. The real data part is weak, and lacks comprehensive discussion, detailed illustrations, and interpretations. The authors should provide a thorough data analysis to convincingly demonstrate to practitioners the validity of the causal conclusions drawn in the paper.

4. From a causal methodological point of view, as highlighted by one referee, the target parameter of interest \theta_{C,1}-\theta_{C,0} is not identified by Theorem 2.1 as \theta_{C,0} is defined differently as specified on page 8. This constitutes a serious error that calls into question the overall validity of the paper.

5. The decomposition in Figure 1 also looks confusing to me. The authors should elucidate each component with clear, straightforward language and provide their respective meanings. Furthermore, a clear explanation of the necessity for P_C and P^*_C, along with their interpretations, would be beneficial.

6. The authors contend that since the disease status A cannot be manipulated, they express the expected potential outcomes as conditional expectations. To enhance readability, it is recommended that the authors adopt mediation terminology, denoting \theta_{O,a}=E(Y(a)) and incorporating the ignorability assumption. Additionally, the inclusion of a directed acyclic graph would more clearly delineate the interrelations among the variables. If there are concerns regarding the manipulability concept, a note could be added to clarify that these

can be considered conditional associations for those skeptical about the manipulability of disease status.

7. The authors use mean framewise displacement, a summary of measure of motion, which is an unsatisfactory simplification. Presumably, this approach is taken because the existing theories apply only to a continuous mediator M. For the paper to meet the standards of JASA ACS, the authors should analyze the full scope of motion data rather than a summary statistic. This approach becomes particularly pertinent when employing cross-fitting machine learning methods that are more effective with multi-dimensional M.

In addition to the referees’ other comments, I have several minor comments here:

1. I agree with one referee that the authors should not remove subjects that are “corrupted by motion."

2. I find the authors' claim in the introduction, that confounder regression fails to mitigate the effects of motion due to non-linearity or unaccounted-for motion effects, to be unconvincing. By employing cross-fitted machine learning methods and considering motion as a confounder, one can indeed adjust for non-linear effects.

3. Upon examining Figure 3, it appears that the methods showcased do not exhibit notable distinctions.