Great job, Maggie!

Just a few comments from me:

Schizophrenia typically has a later onset than adolescence (first psychotic episode, diagnosis and so on) although there are often some earlier symptoms as you correctly write. So perhaps rephrase the first sentence to something more like «Schizophrenia is a complex mental disorder with symptoms that typically originate during adolescence, a critical period…..»

The fatality risk/shorter lifespan is mostly due to high rates of suicide and lifestyle, so rephrase that a bit.

Schizophrenia also has high heritability/genetic component, which is also highlighted in existing research. Even though this is not a focus of your study, consider including that among the potential causes and influences to give a thorough background.

Very nit-picky, but write out «timepoint 1», «timepoint 2» as it can be mistaken for imaging types at first glance (MRI, T1 weighted vs T2 weighted images)

Also agree with Esten and the rephrasing of the research question!

I second Esten; really impressed by how quickly you put together this project description and all the relevant references.

I don’t have many other points than what Esten has already mentioned, but if you need input to the part about ethics/economics, I have some suggestions below:

Ethical considerations and applications: The data need to be kept in a safe place, and we use UiO’s TSD service for this.

Economics / resources needed: Access to TSD (project p23) and computing resources for running machine learning models, provided by the Center for Lifespan Changes in Brain and Cognition.

This is really good Maggie, well done!

Two minor comments from here:

- I would tone down the longitudinal aspect, as I'm not entirely sure how this should be done. Maybe putting point (2) first (I would also call this a predictive analysis instead of a complexity analysis), and saying something like:

"Assess whether longitudinal data from the two first timepoints can increase the accuracy compared to a model based solely on data from a single timepoint".

- I would phrase the research question around complementary predictors instead of mentioning optimality, e.g.

"Understand how different data modalities complement each other towards predicting phenotype ahead of time"

- Esten