Preregistration:

1. Have any data been collected for this study already?
2. What’s the main question being asked or hypothesis being tested in this study?

* What factors influence the “se faire” vs. “être” passive alternation in French?
* In other words: what factors influence speakers’ choice between the “se faire” and the “être” passive construction?
* Operationalized as: the probability that the “se faire” passive construction is used

1. Describe the key dependent variable(s) specifying how they will be measured.

* DV: type of passive construction (“se faire”/“être”)
* Values: 0 = “être” construction; 1 = “se faire” construction
* Annotation criteria: …

1. How many and which conditions will participants be assigned to?

* This is a corpus study, so there will be no participants.
* Factors for the logistic regression:
  + Medium? (informal/formal)
  + Animacy of the subject (inanimate/animate)
  + Person of the subject (...)
  + Number of the subject (…)
  + Gender of the subject? (...)
  + Complément d’agent (absent/present)
  + ...
* Interactions tested: ...

1. Specify exactly which analyses you will conduct to examine the main question/hypothesis.

* Logistic regression
* The model will calculate the probability that a given construction is realized with „se faire“
* The predictor variables will be normalized (log of the odds)
* Random effect: lemma of the verb (to avoid individual effects of the verb)
* p-value 0.05?

1. Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

* Constructions will only be included if:
  + the construction can be realized with both “se faire” and “être” passive
  + the verb is a transitive verb
  + tense of the verb?
* ...

1. How many observations will be collected or what will determine sample size?
2. Anything else you would like to pre-register?