# Supplementary Materials



**Supplementary Figure S1:** Qualitative assessment of the performances of the NLP method for matching text to the standard vocabulary. (**a)** Raw muscle histology report text with detected keywords highlighted in green and red. A red highlight indicates that the keyword is in a negated sentence. **(b)** Table of some highlighted keywords and the details of the match (matching vocabulary ID and terms, position in the raw text, matching n-gram [raw text] and the similarity score of the comparison). Green and red colors correspond to keywords detected as present or present in negated sentence respectively.