Software based on Behavioral economics: a novel effort to raise correlation levels in LVI breast cancer:

    Breast vascular invasion is a high yield factor in breast cancer prognosis, a fact that has been established in different papers; although, its recognition is far from being a simple task. Many authors have proposed different criteria to address its correct identification in H&E slides. Nonetheless, different reviews have pointed out high variability issues among highly trained pathologists.

    The inter-observer variability problem is an issue that could be tackled in different ways, but essentially, it is a categorical problem. For example, let’s imagine a hypothetical situation where,  in order to follow a decision algorithm, you must decide yourself whether an object is black or white; although, the object is grey. This hypothetical situation could be tackled differently according to different strategies:

    a)The observer could decide to look closer, so depending on how many pixels of black he counts,  he could decide if it is black or white.  Although, eventually, he will confront the same struggle (where to set up the cut point off?).

    b) Creation of an “unknown or not sure” category. This could improve the concordance issue; however, this category won’t have any action or consequence associated with the algorithm, so basically, it becomes useless.

    c) Creation of a “grey” category, and allocate a new algorithm action to this category. However, the problem remains unsolved  because the observer will still be struggling between grey/white and grey/black

    This hypothetical situation occurs, the same, in many different settings; although, their results are different for many different reasons. Let’s think about it; if pathologists are able to identify one single neoplastic cell in an H&E slide from a paraffin-embedded block, of course we know the criteria to diagnose  neoplastic cells, although the most probable scenario would be that a huge group of pathologists would differ about whether or not there is a tumoural cell in that slide. However, it is not a big deal, most of the time, to diagnose cancer cells in H&E slides because we are not required to do it in such a demanding way; the most common scenario would be at least thousands to millions of cancer cells to be diagnosed in the paraffin-embedded block, so no correlation problem there. The LVI problem could be summarized, as the former examples suggest, in a quantity problem; we are so close and the borderline is so thin that we struggle to say what is going on.

    In short, this problem will not be solved by adding categories or making a more precise observation ( training).On the contrary,  a different approach could be given by behavioural economics, where the answer to this problem would be to put the observers on the spot, so they could be influenced in order to raise their levels of agreement compared with a determined standard or focus group.

    Behavioural economics is a method of economic analysis that applies psychological insights into human behaviour to explain economic decision-making. As an illustration, Chile had a major issue related to the number of organ donations; its former legislation entitled people to sign up to be an organ donor at a public office if they want to be one. The authorities tackled this issue, amending the law, making the default option to be a donor; unless, the citizen goes to sign up to not be a donor. This single measure raised the levels of donors in the country.

    As an example, levels of agreements could be raised by only fixing a pre-determined option; if the pre-determined option establishes that there is a vascular invasion, it is expected that the number of pathologists that will agree with that observation rises. This is an example of using inertia and the power of default arrangements in order to increase levels of agreement.

    Even though the agreement problem could be addressed this way, an issue of randomness comes out; how do I set the predetermined option throughout different cases? (specimens and slides). Thereupon an experimental equation was made to set up the pre-test probability of having LVI in the slide. This equation considers the following variables: age, tumour size, Nottingham grade, and the number of LVI images seen on the slide. Also, it is based on the data collected by the following studies: The Prognostic Significance of Lymphovascular Invasion in Invasive Breast Carcinoma (<https://www.ncbi.nlm.nih.gov/pubmed/22180017>) and Diagnostic concordance of reporting lymphovascular invasion in breast cancer (<https://www.ncbi.nlm.nih.gov/pubmed/29599396>).

    The project involves creating prospective research, where we could provide the subjects with some slides, so the pathologists would give us their impression of what is going on. In order to do that, a small software program should be designed in order to explore the utility of behavioural economics on the detection of vascular invasion in breast cancer.

    First, it will be mandatory to gather a few virtual slides of representative cases.  Second,  a questionnaire should be designed and probably other questions should be addressed too (such as cancer type, DCIS and others); the Nottingham EQA scheme could be a very appropriate setting for this experimental software. Third, a probability of LVI will be provided to the subject; this probability would be obtained by the equation.

    Let’s explain in more detail. The observers will have to choose between presence or absence of LVI. They will be split into two groups, control group and intervention group. The intervention will consist of introducing,  to the observer,  a visual "guide" that could be one or more of the followings ideas:

A slide agreement ratio and an overall agreement ratio, based on kappa coefficient, would compare the observers in real time or in a different time (The timing could change the result, to review in more detail please refer to game theory static and dynamic games).

The probability of LVI in that particular setting (already discussed).

Both the ratios and the probability of LVI could be highlighted by different colours depending on whether the rate of agreement is below than substantial agreement (Kappa=0.61-0.80) or depending on the theoretical probability of LVI.

    In addition, a post-answer nudge could be provided in the case that disagreement is noticed:

A warning text could be displayed in red capital letters and a password window could pop up when the provided answer has a Kappa below 0.61. This should discourage audacious answers.

Relying on the data gathered from other observers in the experiment, the software could suggest a more concordant answer; therefore, the subject will have to choose if he prefers his own answer or the software’s answer.

     The primary goal of this project is to compare the Kappa ratio between both groups. If a statistical significance is noticed, we could repeat the same experiment on a bigger scale and, later on, the do  more research about whether this improved Kappa ratio maintains its prognosis significance.