Instructions for Reporting the *Confidence Rating*

The *Confidence Rating* is a rating given to each cell that reflects your level of confidence that the cell is one of the "target" cell types.

The *Confidence Rating* we are asking you to provide is not part of a standard clinical report and you may not be trained to provide such a *value.* The purpose of reporting a *Confidence Rating* is to provide more information than just the binary classification: positive (the cell is one of the target types) vs. negative (the cell is not one of the target types). The *Confidence Rating* is meant to provide more robust, quantitative, and statistically powerful data for evaluating the imaging technology in the non-clinical setting.

The scale for the *Confidence Rating* arises from the following question:

Given two cells, which one are you more confident is calling the target cell type?

For example, what if you called two cells as targets? Could you declare one cell as a target with more certainty? On the other hand, what if you decided that two cells were not targets? Could you rank one as more likely not a target compared to the other? More generally, could you take a stack of images of cells and lay them out on a conference table, sorting them according to your confidence that they are targets (allowing for some ties)?

Assuming you can compare cells, you have a latent (hidden) confidence rating. The *Confidence Rating* is this rating. Determining and expressing this rating is the challenge. The purpose of this document is to provide some descriptions, anchors, and tools to help you report a *Confidence Rating*.

Meaning of *Confidence Rating*

We will use the integers from 0-100 for the *Confidence Rating.*

- 0 is the lowest *Confidence Rating*. Use this when you are absolutely certain that the cell is not one of the target types.
- 100 is the highest *Confidence Rating.* Use this when you are absolutely certain that the cell is one of the target types.

To help anchor your scores, we define some boundaries/thresholds.

- 0 is a red flag. Use this score when a location doesn't remotely have a candidate cell (no tissue, no cell, inappropriate cell).
- 25 is a point of hesitation. If you immediately and instinctively reject a cell as a target, score it 25 or less. If you deliberate, consider individual features, and weigh the odds before resigning to a decision that a cell is not a target, score it between 25 and 50.
- 50 is the point of equivocation. Numbers 50 and below this should be given to cells that you would ultimately *NOT* call targets (NEGATIVE CALL). Numbers above 50 should be given to cells that you would ultimately call targets (POSITIVE CALL).

- 75 is another point of hesitation. If you immediately and instinctively want to call a cell a target, score it 75 and above. If you deliberate, consider individual features, and sum up all the evidence in determining that a cell is a target, score it between 50 and 75.

Being Quantitative

Over time within an experiment, you should attempt to make your *Confidence Ratings* as quantitative as possible in terms of relative comparisons with previous scores: you should be more certain that a cell given a *Confidence Rating* of 45 is a target compared to a cell given a *Confidence Rating* of 35.

If you use a *Confidence Rating* rating more than once, you are indicating that there are no features indicating one cell is more likely than another to be a target.

You should try and space out your *Confidence Rating* ratings to allow for a new cell that has a *Confidence Rating* in between. For example, your first reaction to a cell is to score it 20. However, you recall another cell that you scored 20, and you are less confident that the current cell is *not* a target (more confident that the current cell is a target). Instead of scoring the current cell 21, perhaps score the current cell 23 or even 25.

Being quantitative is not easy. Do your best. If you are only comfortable using 10, 20, 30, … That's fine.

Good Luck!