# Artificial Intelligence Assisted Assessment of HER2 IHC Using Semi- and Fully-Automated Workflows: Rapid Validation and Implementation as a Quality Assurance and Risk Mitigation Check Point

Ursula Green[1], Jenifer Caldara[2], Danielle Pirain[2], Regan Baird, PhD[2], Jochen K. Lennerz, MD, PhD[1,3]
[1] Massachusetts General Hospital, Boston, Massachusetts, [2] Visiopharm Corporation, Westminster, Colorado, [3] Harvard Medical School, Boston, Massachusetts

## Background:

Human Epidermal Growth Factor Receptor 2 (HER2) is a standard and critical biomarker in the staging, diagnosis, and management of breast cancer. Proper classification of HER2 status in stained tissue sections is critical to precision pathology and personalized medicine for these cancer patients. Borderline cases require additional timely and costly assessments, including reflex and in situ hybridization (ISH) testing. Visiopharm's HER2 algorithm (APP) has a proven track-record of assisting pathologists by expediting review time during staging and identifying borderline cases which are immunohistochemistry (IHC) positive, ISH normal as having a HER2 score of 1+. The identification of this population has been traditionally challenging through manual IHC interpretation, and AI can provide a crucial quality control and risk mitigation checkpoint.

In this study, we suggest that the HER2 APP may be used as a tool to confirm manual sign-out assessments, or to flag specimen which are re-assigned to a new classification by the APP. This APP would be applied following the manual sign-out as a method of mitigating risk on challenging cases through the labeling of non-concordant reads during this quality assurance check point. Here we additionally validate the use of the HER2 APP using retrospective analysis on known patient samples with both IHC and FISH manual assessments across multiple hospitals to demonstrate the use-case for the Center of Integrated Diagnosis (CID).

## Methods:

58 random patient samples were stained for HER2 IHC at one of 5 hospitals according to their standard staining procedures (Figure 1). Samples were assessed by both standard pathology and FISH. Samples were then digitized at the Massachusetts General Hospital (MGH) CID on a Motic Scanner and tumor areas analyzed by the HER2 (Visiopharm) APP. Analysis was performed by A: a semi-automated workflow where areas of tumor selection were guided by pathologist inking prior to imaging; and B: a fully-automated workflow where the tissue was compartmentalized by outlining areas of tumor using Artificial Intelligence (AI) prior to HER2 analysis (Figure 2). Scores from both methods were then compared to the manual pathology IHC and FISH interpretations.

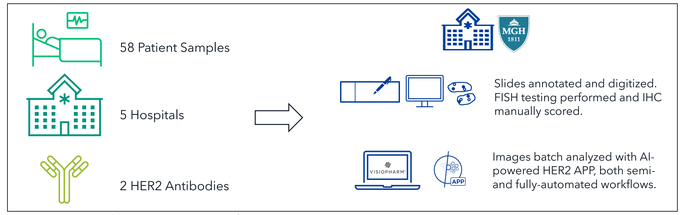


Figure 1: Visual representation of specimen movement and process during study

## Methods Cont':

**Semi-Automated:** Regions for analysis are placed by the pathologist

**Fully-Automated**, Whole-Tissue: Regions for analysis are automatically placed and later confirmed by a pathologist

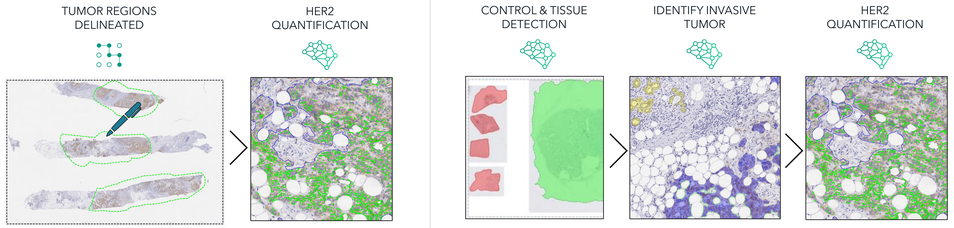


Figure 2: Visual representation of the process followed for the Semi-Automated (left) and Fully-Automated (right) workflows to identify tumor regions and score HER2 IHC staining

## Results:

Table 1: The correlation between the Visiopharm HER2 APP score and the IHC manual score, as well as the FISH results for IHC 2+ specimens, relative to the two tested workflows.

- Concordant
- Neural Impact: Discordant IHC, No Impact on Treatment Strategy (Minor discordance)
- Potential Positive Impact: Discordant IHC, FISH Investigated. Possible Impact on Treatment Strategy
- Negative Impact: Discordant IHC, Deleterious Impact on Treatment Strategy (Major discordance)

| HER2 APP Score | Manual IHC Score 1+ | Manual IHC Score 2+ | Manual IHC Score 3+ | FISH Results on IHC 2+ Specimens: Normal | FISH Results on IHC 2+ Specimens: Abnormal | |
|---|---|---|---|---|---|---|
| 1+ | 6 | 19 | 0 | 17 | 2 | **Semi-Automated Workflow** |
| 2+ | 2 | 8 | 0 | 5 | 3 | |
| 3+ | 1 | 3 | 19 | 3 | 0 | |
| 1+ | 7 | 22 | 0 | 21 | 1 | **Fully-Automated Workflow** |
| 2+ | 2 | 6 | 0 | 3 | 3 | |
| 3+ | 0 | 2 | 19 | 1 | 1 | |

## Discussion/Conclusion:

17 and 21 cases were appropriately reassigned from a 2+ to a 1+ score by the HER2 APP using the semi- and fully-automated workflows, respectively (Table 1). The labeling of these cases with a HER2 score that was different from the IHC score given during sign-out demonstrates real-world examples of cases which would be flagged for an additional examination by the pathologist to confirm the proper classification. Furthermore, the HER2 APP served as an additional, independent quality-control checkpoint by confirming the IHC status on a case and may occur before, during, or after the FISH results are obtained. By alerting that a given sample may demonstrate IHC inconsistency, the pathologist is able to (a) revisit the sample to confirm the IHC results, and (b) interpret the FISH results using an independent corresponding IHC result (e.g., same region of interest).

It was also seen that the AI was not influenced by the different hospitals' utilization of various HER2 clones and staining protocols. Cases showing a major discordance using the fully-automated workflow were primarily due to the accidental detection of excision and staining artifacts by the AI. These artifacts would easily be corrected by an observer, confirming that the AI is a tool to be used in conjunction with the observer for the highest level of precision medicine. Inclusion of these artifacts in the AI training set will improve performance of both workflows. Additionally, our study demonstrates that the HER2 APP may be validated and implemented in pathology laboratories within 48 hours for the purposes of quality control and risk mitigation. This may also be crucial in recognizing minimally positive specimens which may qualify for HER2-low status and treatment strategy, different from traditional HER2-positive therapy.

## Summary:

**Both approaches yielded results which validated Visiopharm's HER2 APP for use in the MGH laboratory's routine workflow.**

- Flagging cases where the HER2 APP score is non-concordant with the IHC sign-out score may serve as a risk mitigation check point for pathologist verification
- The ability to rapidly validate and implement such quality assurance check points may be critical when novel tests emerge, such as HER2-low assessments