

# Validation Plan

## Name of the device:

HippoQuant.AI

## Intended Use:

To aid radiologists in identifying the hippocampus in MRI head images and quantifying its volume.

## Training dataset information:

The training dataset consists of MRI head images with a focus on the hippocampus region. Labels have been manually created by experienced radiologists. Several factors, such as the expertise level of the annotating radiologist and time constraints, can influence the accuracy of the manual labeling process.

## Patient Population Description for Validation Dataset:

Patient characteristics:

- patient age: >20
- patient gender: male or female

Image characteristics:

- imaging modality: MRI
- body part examined: head

Patient condition:

- Patients may be in any condition, including neurodegenerative and other diseases that may impact hippocampal volume

## Ground Truth Acquisition Methodology:

To establish a robust ground truth, three radiologists will independently label the hippocampus in the MRI images. The labeled regions will be designated as either anterior or posterior. A consensus mask will be generated by averaging the individual masks, with pixel intensities divided by 3. A threshold of 0.5 will be applied to classify each pixel as either hippocampus or background. This methodology reduces the likelihood of errors and increases the accuracy of the resulting mask, especially along the hippocampus boundaries.

## Algorithm Performance:

The performance of the algorithm will be evaluated using the following metrics:

- Dice score
- Jaccard score
- Sensitivity
- Specificity

Upon obtaining the ground truth mask, the DICOM image and the mask will be input into HippoQuant.AI. The algorithm will generate its own mask, which will then be compared to the ground truth mask to

calculate the performance metrics. While the Jaccard score, Sensitivity, and Specificity serve as supplementary metrics to support the evaluation, the primary metric, Dice score, should be  $\geq 0.9$ . In addition to the numerical results, the predicted mask outline will be superimposed onto the image to facilitate a visual assessment of the algorithm's accuracy.

The primary objective of the algorithm is to determine hippocampal volume. Once the mask has been generated, HippoQuant.AI will calculate the anterior, posterior, and total hippocampal volumes. Radiologists will verify the results and compare them to existing literature (e.g., the HippoFit calculator).