# Validation Plan

* What is the intended use of the product?

The intended use of the product is to help clinicians measure the volume of the hippocampus in a faster and more consistent manner. It is an AI system that segments and measures the hippocampal volume and integrates into the clinicians workflow.

* How was the training data collected?

We are using the "Hippocampus" dataset from the [Medical Decathlon competition](http://medicaldecathlon.com/). This dataset is stored as a collection of NIFTI files, with one file per volume, and one file per corresponding segmentation mask. The original images here are T2 MRI scans of the full brain. As noted, in this dataset we are using cropped volumes where only the region around the hippocampus has been cut out.

Through a multi-institutional effort, researchers generated a large, curated dataset representative of several highly variable segmentation tasks. The hippocampus dataset consisted of MRI acquired of the right hippocampus in 90 healthy adults and 105 adults with a non-affective psychotic disorder. All subjects were free from significant medical or neurological illness, head injury, and active substance use or dependence. Structural images were acquired with a 3D T1-weighted MPRAGE sequence (TI/TR/TE, 860/8.0/3.7 ms; 170 sagittal slices; voxel size, 1.0 mm3 ). All images were collected on a Philips Achieva scanner

* How did you label your training data?

All data has been labeled and verified by an expert human rater, and with the best effort to mimic the accuracy required for clinical use. These labels were used as the ground truth.

* How was the training performance of the algorithm measured and how is the real-world performance going to be estimated?

The training performance was measured using Dice and Jaquard Score. Real world performance will be estimated by comparing to brain volume percentiles from real world datasets and ensuring the value is in the appropriate ages. Given that we have no budget or contraints we could hold a large scale clinical trial, capture the hippocampal scans of a diverse population and have several expert clinicians compare the performance of the algorithm with that of their own expert measurements. This would guarantee that the technical performance of the device was satisfactory. Additionally, the representatitve users should use the device on a representative patient population under the expected conditions of use to validate the clinical performance of the device.

* What data can the algorithm operate on?

The data would perform best on MRI scans of the brain that are cropped so that only the region around the hippocampus is cut out. The patients should be adults and healthy or with a non-affective psychotic disorder and free from significant medical or neurological illness and head injury. Ideally images would be collected on a Phillips Achieva Scanner and would be saggital slices.