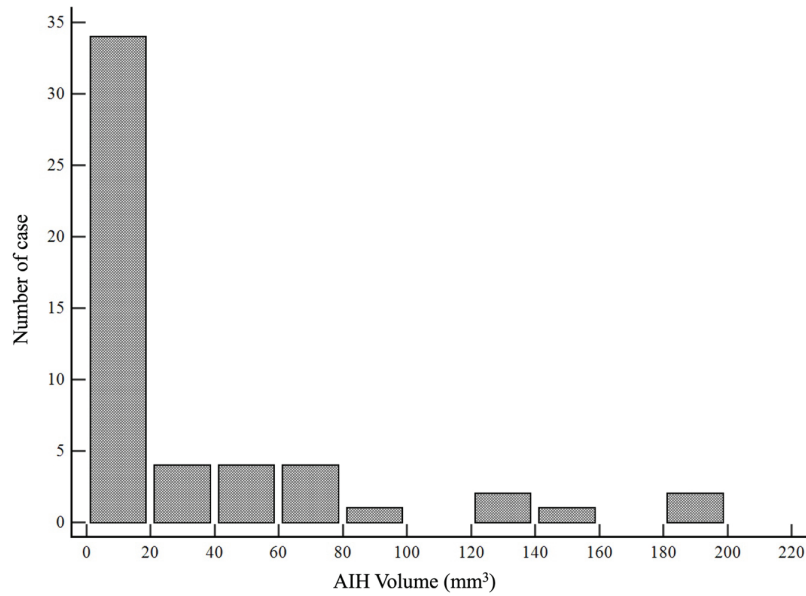
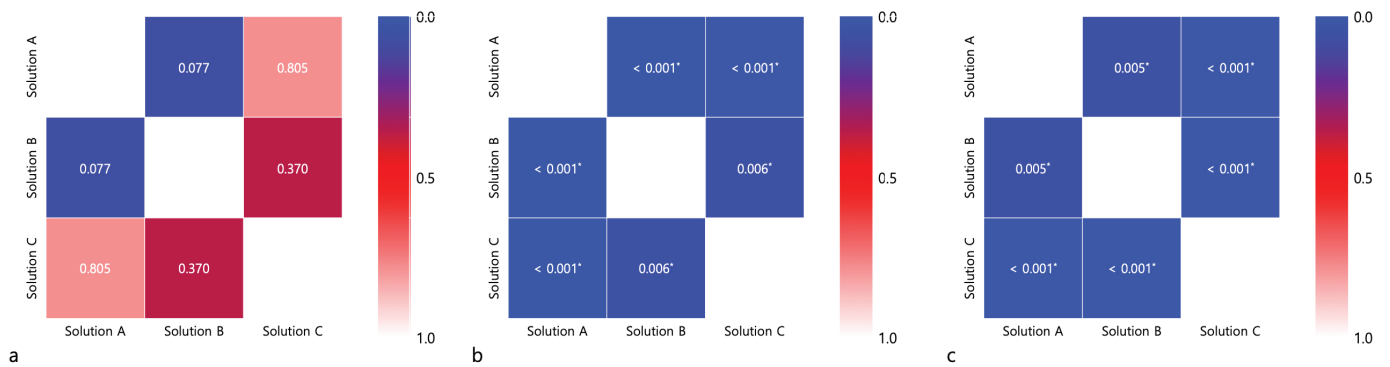
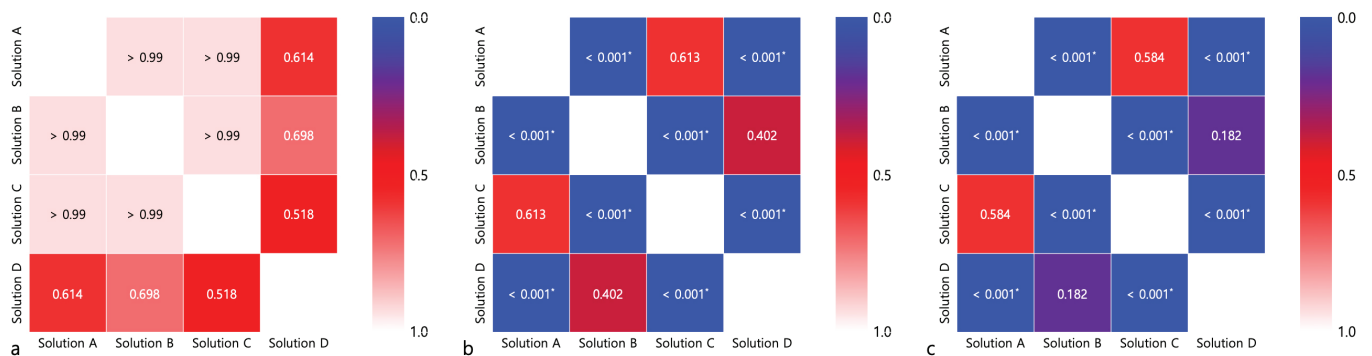




Supplementary Figure 1. Histogram of AIH volume measured by neuroradiologists. AIH, acute intracranial hemorrhage.



Supplementary Figure 2. This figure illustrates heatmap plot of P values in detection performance of solution A, B, and C based on probability score. (a) Pairwise comparisons of AUROC among the four solutions showed no statistically significant differences (P values > 0.05). (b) The AUPRC and (c) Brier score among solution A, B, and C were significantly different (P values < 0.01). Asterisks (*) depict the P values with statistical significance. AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve.



Supplementary Figure 3. This figure illustrates heatmap plot of P values in detection performance of solution A, B, C, and D based on binary classification. (a) Pairwise comparisons of sensitivity among the four solutions showed no statistically significant differences (P values > 0.05). (b) Differences of specificity were statistically significant, except for the pairwise comparisons between solution A and C, and solution B and D. (c) In precision and F1 score, differences were statistically significant, except for the pairwise comparisons between solution A and C, and solution B and D. Asterisks (*) depict the P values with statistical significance.