

Using a dataset of 39 known experimentally obtained disordered proteins/domains, they found 77% of the proteins were correctly predicted by FoldIndex, 72% by PONDR VL-XT and 20% by GlobPlot respectively. We used the same dataset of 39 unstructured proteins for disorder prediction by IUPRED, DisEMBL and RONN to check the accuracy of predictions by these tools. RONN and IUPRED predicted 58.33% correctly and DisEMBL only 19.4%. Further, a dataset of 46 proteins was compiled from a search of HD proteins from the NCBI's Entrez Gene database by giving the query "Huntington disease Homo sapiens". All of these proteins were checked for disorderedness using all the six tools stated above. Of these, 34 and 36 proteins were predicted as unstructured by FoldIndex and PONDR VI-XT respectively. Globplot, DisEMBL, IUPRED and RONN predicted only 18, 19, 14, and 18 of the total 46 proteins as disordered, respectively. Owing to these findings, GlobPlot, DisEMBL, IUPRED and RONN, having lesser accuracy level, were not used any further in our analysis. Hence, FoldIndex, with a reported accuracy rate of 77% was used all through, for all the disease and control datasets.