Benchmarking and Assessment for Multiple Imputation