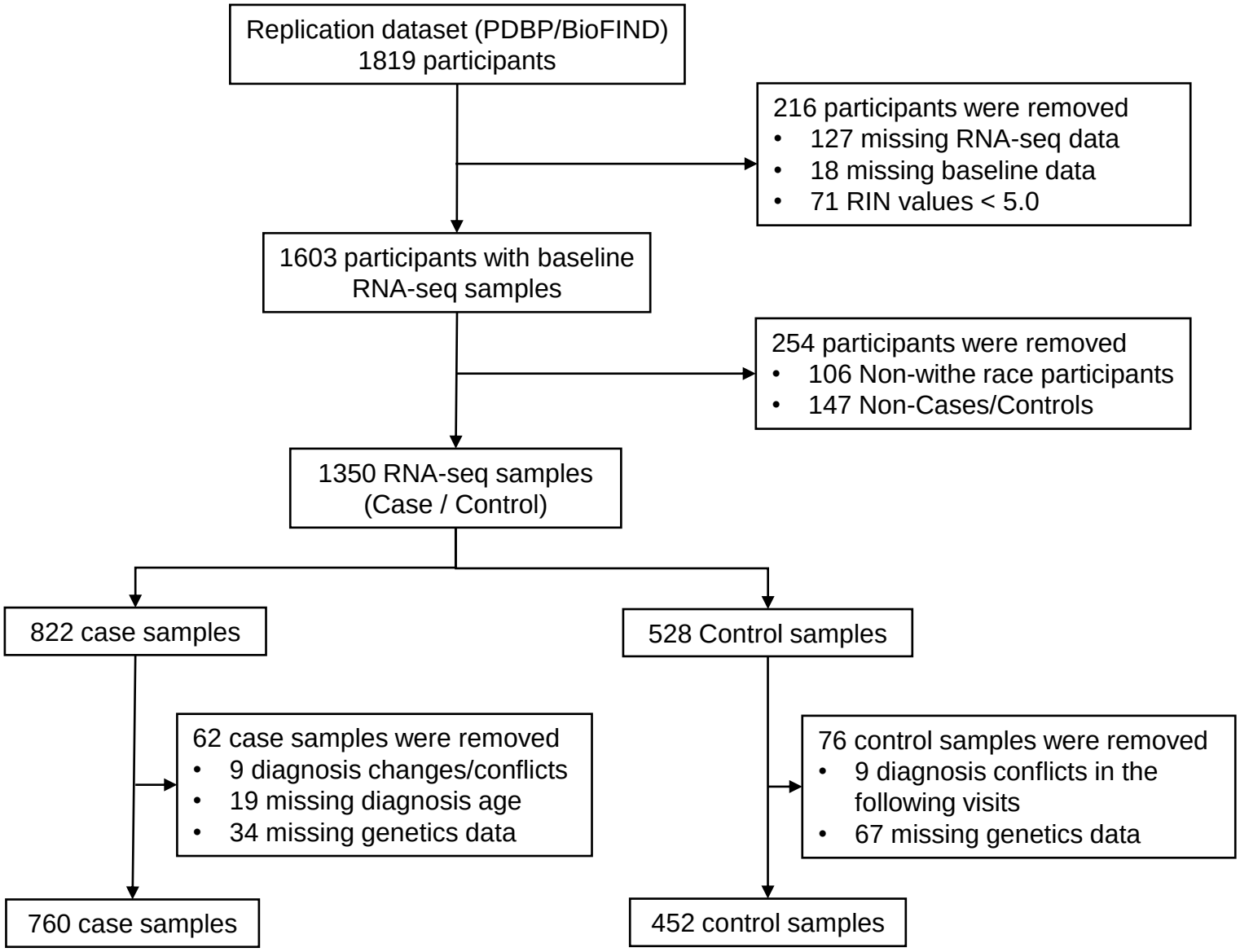
Replication dataset (PDBP/BioFIND)
1819 participants
216 participants were removed
• 127 missing RNA-seq data
• 18 missing baseline data
• 71 RIN values < 5.0
1603 participants with baseline RNA-seq samples
254 participants were removed
• 106 Non-withe race participants
• 147 Non-Cases/Controls
1350 RNA-seq samples (Case / Control)
822 case samples
528 Control samples
62 case samples were removed
• 9 diagnosis changes/conflicts
• 19 missing diagnosis age
• 34 missing genetics data
76 control samples were removed
• 9 diagnosis conflicts in the following visits
• 67 missing genetics data
760 case samples
452 control samples