**Figure S2. Data partitioning in Random Forests**

**Original dataset**

n=**6,522**

**Validation Partition**

n=1.956

**Derivation Partition**

n=4.566

**Out of Bag**

(1/3 of derivation set)

**Train data**

(2/3 of derivation) **set)**