PCA: MeSH Term Clusters within Estonian Biobank
(10 semantic clusters from TF-IDF analysis)

C2 (n=35)

C9 (n=66)

C0 (n=70)

C4 (n=82)

C6 (n=65)

C7 (n=133)

C1 (n=27)

C5 (n=62)

C3 (n=22)

C8 (n=25)

PC2 (16.7% variance)

PC1 (20.2% variance)
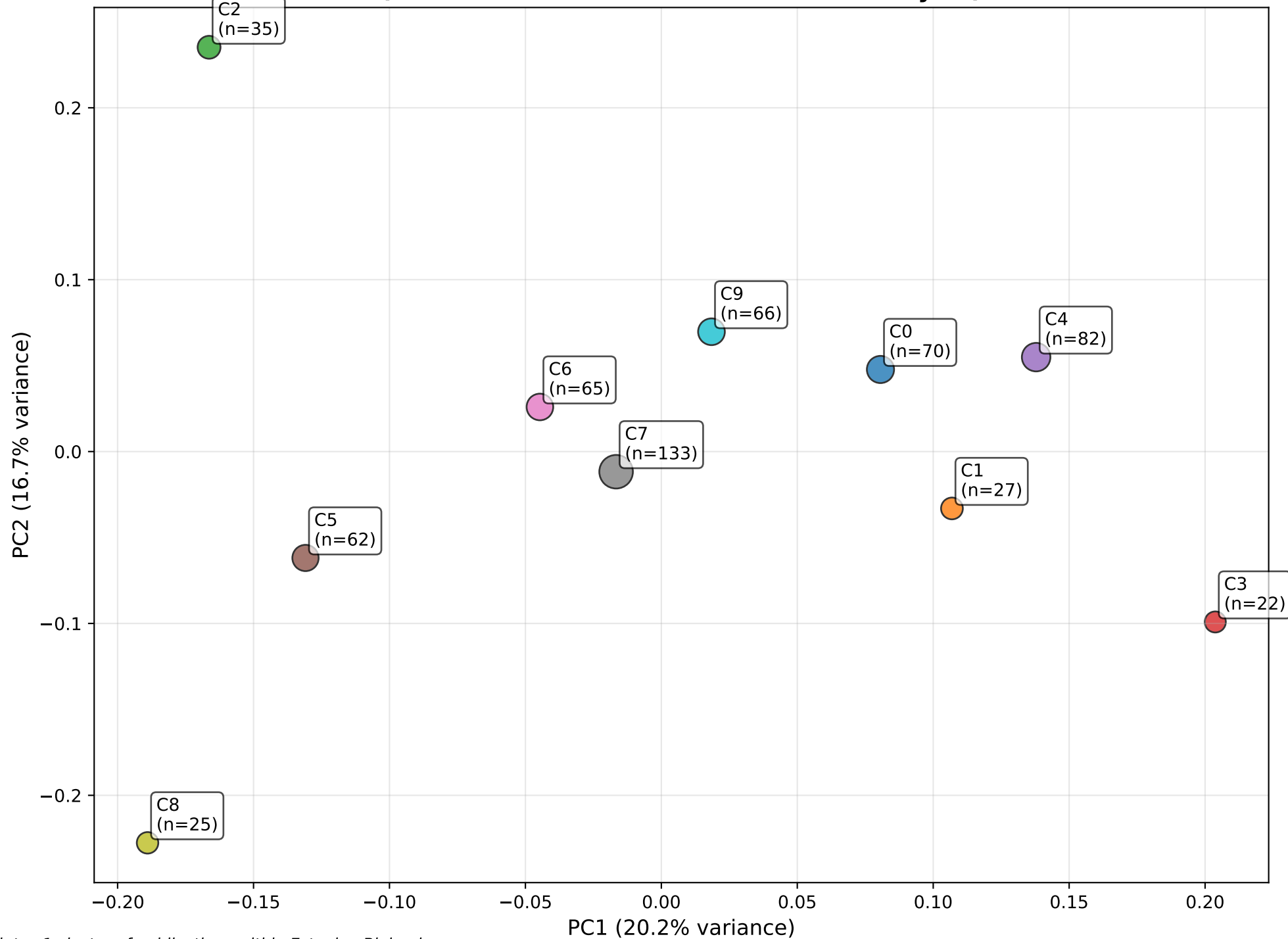
Each point = 1 cluster of publications within Estonian Biobank
Point size ∝ number of publications in cluster
Distance = semantic similarity of MeSH terms