# Supplementary Table – Cluster Descriptions

|  |  |  |  |
| --- | --- | --- | --- |
| Cluster | # Rows | Top 5 ICD-10 Codes (frequency) | Dominant Semantic Terms (top 5) |
| F0 (female) | 869 | 1. F17210 (310) 2. F1210 (113) 3. F603 (74) 4. F17200 (45) 5. F1290 (42) | • nicotine\_dependence,\_cigarettes,\_uncomplicated • cannabis\_abuse,\_uncomplicated • borderline\_personality\_disorder • personality\_disorder,\_unspecified • conduct\_disorder,\_unspecified |
| F1 (female) | 742 | 1. F17210 (425) 2. F17200 (127) 3. F1210 (53) 4. K3580 (52) 5. J3501 (34) | • nicotine\_dependence,\_cigarettes,\_uncomplicated • nicotine\_dependence,\_unspecified,\_uncomplicated • unspecified\_acute\_appendicitis • chronic\_tonsillitis • cannabis\_abuse,\_uncomplicated |
| F2 (female) | 713 | 1. F17210 (158) 2. O99334 (119) 3. F1210 (69) 4. Z370 (65) 5. F603 (38) | • smoking\_(tobacco)\_complicating\_childbirth • nicotine\_dependence,\_cigarettes,\_uncomplicated • single\_live\_birth • cannabis\_abuse,\_uncomplicated • post-term\_pregnancy |