### DESIGN ELEMENTS

| Study characteristics and considerations | Data source #6 | Data source #7 | Data source #11 |
|------------------------------------------|----------------|----------------|-----------------|
| **Study population**                    | 5              | 5              | 5               |
|   • More than 5 million population (targeted 371 infants with in utero exposure to varenicline) |                |                |                 |
| **Treatment/exposure and comparator group(s)** | 5              | 5              | 4               |
|   • Maternal prescription data linked with birth data |                |                |                 |
|   • In utero exposure to maternal smoking status |                |                |                 |
| **Primary outcomes**                    | 5              | 5              | 4               |
|   • Mother — infant linkage              |                |                |                 |
|   • Birth data linked to infant inpatient/outpatient diagnosis data |                |                |                 |
| **Secondary outcomes**                  | 5              | 5              | 4               |
|   • Birth data linked to infant inpatient/outpatient diagnosis data |                |                |                 |
|   • Birth data linked to infant causes of death data |                |                |                 |
| **Length of follow-up and data recency** | 5              | 5              | 5               |
|   • 1 year after birth minimum           |                |                |                 |
| **Confounding variables**                | 5              | 5              | 4               |
|   • Pregnancy data linked to maternal prescription data |                |                |                 |
|   • Pregnancy data linked to maternal inpatient/outpatient diagnosis data |                |                |                 |

### DATA ACCESS CONSIDERATIONS

| Timeline                     | Moderate | Fast | Moderate |
|------------------------------|----------|------|----------|
|   • Time to fully executed contract |          |      |          |
|   • Time to data access      |          |      |          |
|   • Time to analyze         |          |      |          |

| Contracting logistics       | Medium | Medium | Low |
|-----------------------------|--------|--------|-----|
|   • Cost                    |        |        |     |

### FINAL DATA SOURCE SELECTION

✓✓

---

**LEGEND**

5 = Many/nearly all data requirements met
4 = Several data requirements met
3 = Likely that several data requirements are met but requires further investigation
2 = Some data requirements met or unable to assess at this time
1 = Data requirements not met

---

Figure S2.
Case Example 2: SPIFD Step 3 heatmap retrospectively applied to varenicline pregnancy study