**Accompanying Files:**

* **collider.html**  
  This file contains all the code and output from a simulation illustrating **collider–stratification bias**. It explores how smoking can appear protective for repeat revascularizations, even though it is a known risk factor for initial revascularization. Simply open collider.html in any browser to view the results.
* **collider.qmd**  
  This is the Quarto file that generated collider.html. You can run it in the R ecosystem to reproduce the simulation and outputs.

**Overview of the Simulation:**

* We simulate a deterministic threshold rule to decide whether an individual receives an initial percutaneous coronary intervention (PCI). This rule is chosen **purely for illustrative purposes**—it forces a large imbalance in the unmeasured risk factor U between smokers and non‐smokers. In real-world settings, selection would more likely be modeled **probabilistically** (e.g., via logistic regression).
* Despite smoking being set as a harmful factor (log-odds = 0.3, i.e. an odds ratio ≈ exp (0.3) ≈ 1.35), **conditioning on PCI** can make smoking look protective for repeat revascularizations. This occurs because non‐smokers must have a higher U to “qualify” for PCI, driving up their subsequent risk of repeat PCI and thereby making smoking look protective

**Key Takeaway:**

This demonstration underscores how **conditioning on a collider** (in this case, PCI) can induce a spurious association (or even reverse the true direction of an effect) when the collider is influenced by both the exposure (smoking) and an unmeasured risk factor U. Consequently, one might wrongly conclude that smoking is protective for repeat procedures—even though the underlying data‐generating mechanism in this simulation has stipulated that smoking is harmful.