RoA: visual analytics support for deconfounded causal inference in observational studies

Supplementary clinical use case

We discuss a supplementary clinical use case to elaborate on the various scenarios supported by RoA (see Figure 1) in a step-by-step manner. As a start, we show what the data of a Randomized Controlled Trial (RCT) look like in RoA to highlight the connection between an RCT and an observational study (OBS), shown in Figure 2. Next, we walk through the various other scenarios, for which we refer to the labels shown in Figure 1.

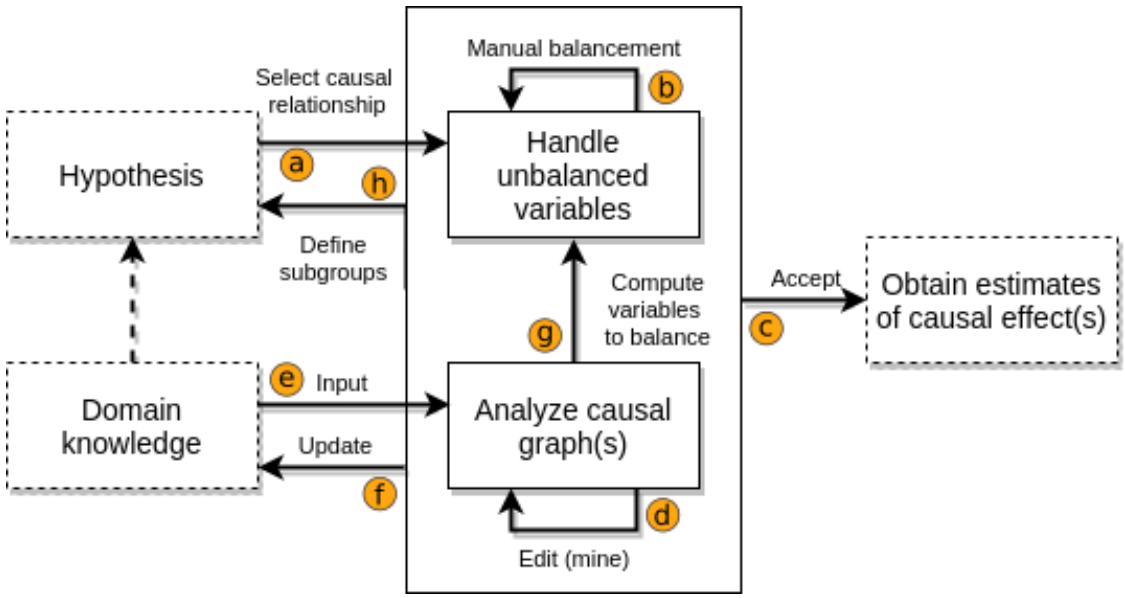


Figure 1: Workflow supported by RoA.

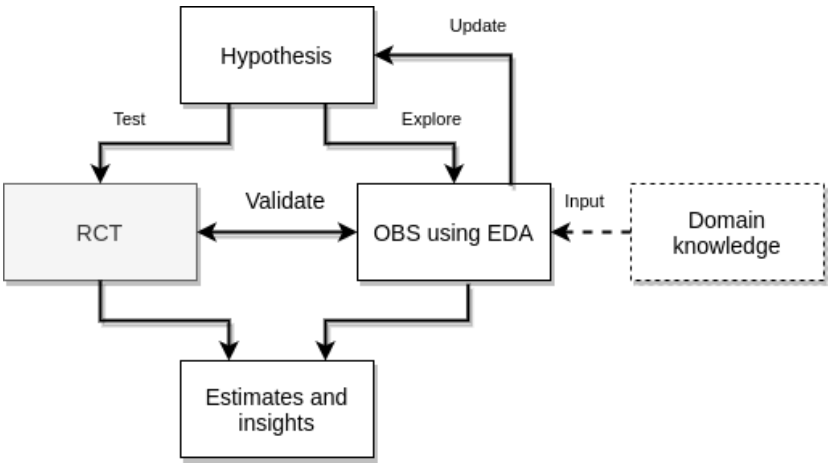


Figure 2: The operational context of the Randomized Controlled Trial (RCT) and the Observational Study (OBS) augmented with exploratory data analysis.

Hypothesis

The dataset used for this clinical use case is based on the so-called FAME study[1] where percutaneous coronary intervention (PCI) was randomized to anatomy- or physiology-guided treatment in 1005 patients. We used a selection of the available variables (shown in Table 1). Furthermore, we have adopted a slightly modified version of the hypothesis from the FAME study: the total length of stents placed depends on the type of PCI treatment, which is either anatomy-guided or physiology-guided. In the former case, arteries with over 50% narrowing are stented, while in the latter case, only the narrowings where a significant drop in blood pressure is measured, are stented. Therefore, our causal relationship of interest is $PCI \rightarrow stentlength$.

The total length of the stents placed is related to the extent of the disease in the coronary arteries, which is called *Coronary Artery Disease (CAD)*. This degenerative disease results in the formation of narrowings in the coronary arteries, which limit the blood flow to the heart to a certain extent. Some of these narrowings limit the blood flow causing reduced presence of oxygen and nutrients being provided to the heart muscle, which causes complaints of chest pain or even a heart attack (myocardial infarction).

Randomized Controlled Trial

Because the *PCI* treatment variable was randomized, we can immediately estimate its effect on the variable *stenthlength*. This scenario comprises steps (a), (b), and (c) shown in Figure 1. We start with selecting our relationship of interest using the case panel and effect panel, as shown in Figure 3. We also select ATE as our estimand to better conform to the FAME study, which implies that our treatment effect estimation captures the extent to which the average total stent length has changed in the general population, depending on the treatment. Next, we observe that all variables are indeed already balanced due to the prior randomization in the balance panel (please note that in this case, the "adjusted" and unadjusted values coincide in the balance panel). Consequently, we can directly obtain the effect estimation from the effect panel: a statistically significant reduction of 13.03 mm in total stent length for physiology-guided PCI compared to anatomy-guided PCI (Glass's delta: -0.0324).

Subgroup analysis

Diabetes patients are of particular interest since this disease is associated with a different expression of coronary artery disease (the presence of smaller vessels) and may therefore have poorer outcomes. To investigate this, we create two subgroups defined by diabetes using the group panel and rerun the analysis in RoA, as shown in Figure 4. This action corresponds to step (h) in Figure 1. We can conclude that the treatment effects within these two subgroups are highly similar to those of the entire study population and that the non-diabetes patients have the least total stent length on average.

[1] Tonino PA, De Bruyne B, Pijls NH, et al. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. N Engl J Med. 2009;360(3):213-224. doi:10.1056/NEJMoa0807611

Variable	Type	Description
pci	binary	treatment of the patient: either anatomy-guided or physiology-guided
male	binary	whether the patient is a male
ccs	categorical	patient complaint class of chest pain: a higher class implies more complaints
unstable_angina	binary	whether the patient has complaints at rest
diabetes	binary	whether the patient has diabetes
smoking	binary	whether the patient smokes
hypertension	binary	whether the patient has hypertension
hyperchol	binary	whether the patient has hypercholesterolemia (high cholesterol)
ef	continuous	the patient's measure of the function of the heart (ejection fraction): the lower the number, the worse the function
asa	binary	whether asperin was given (in case of of high(er) risk profile)
clopidogrel	binary	whether clopidogrel was given: (in case the patient already has a stent, and thus CAD)
beta_blockers	binary	whether beta blockers were given (in case of a high(er) risk profile)
ca_anta	binary	whether calcium antagonist were given (in case of a high(er) risk profile in relation with high blood pressure)
ace_inh	binary	whether ace-inhibitors were given (in case of a high(er) risk profile in relation with high blood pressure)
nitrates	binary	whether nitrates were given (in case of chest pain and thus often CAD)
statins	binary	whether statins were given for counteracting high cholesterol
diuretics	binary	whether diuretics were given to reduce fluid overload in the body, often in patients with low ejection fraction (ef)
insulin	binary	whether insulin was given for countering diabetes
days_censory	continuous	number of days until the event occurred or became censored
death	binary	indicates whether death of the patient occurred after starting the study
mace	binary	indicates a major adverse cardiac event (combination of death and myocardial infarction)
mace_days	continuous	number of days until mace event
stentlength	continuous	the total length of stents (mm) placed after treatment choice (either anatomy-guided or physiology-guided)
family_history	binary	whether coronary artery disease is part of the family history
perivasc_dis	binary	whether the patient has perivascular disease

Table 1: List of selected variables from the FAME study.

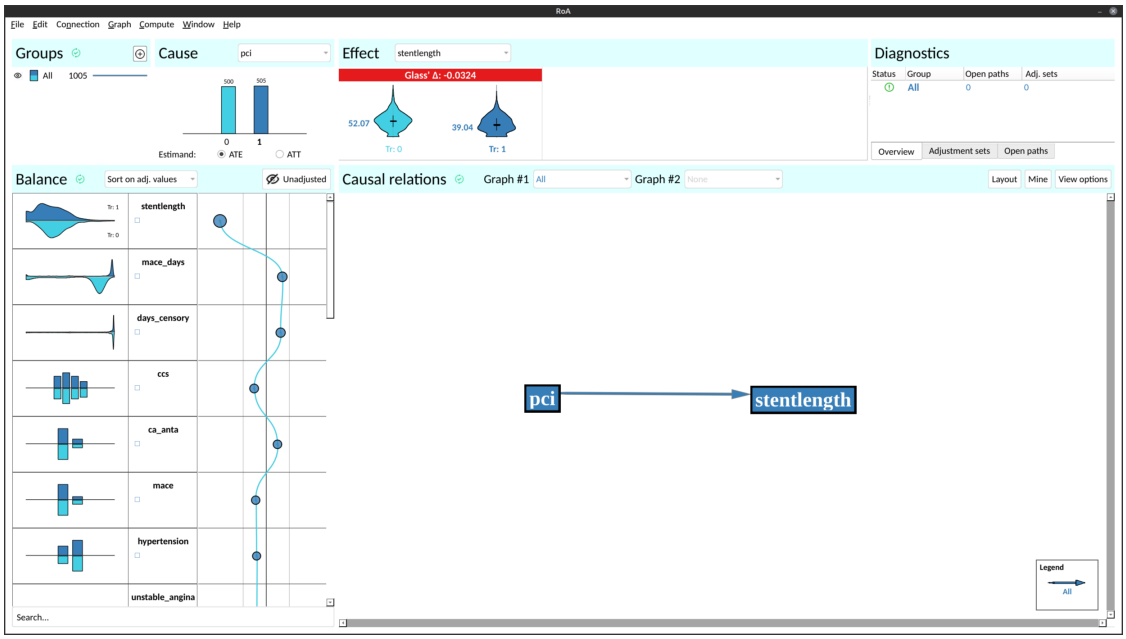


Figure 3: The initial setup for the randomized treatment case. The balance panel shows that the SMD values of the covariates are all within the acceptable range. Note that the adjusted and unadjusted SMD values coincide in this case because no covariates have been adjusted in this case.

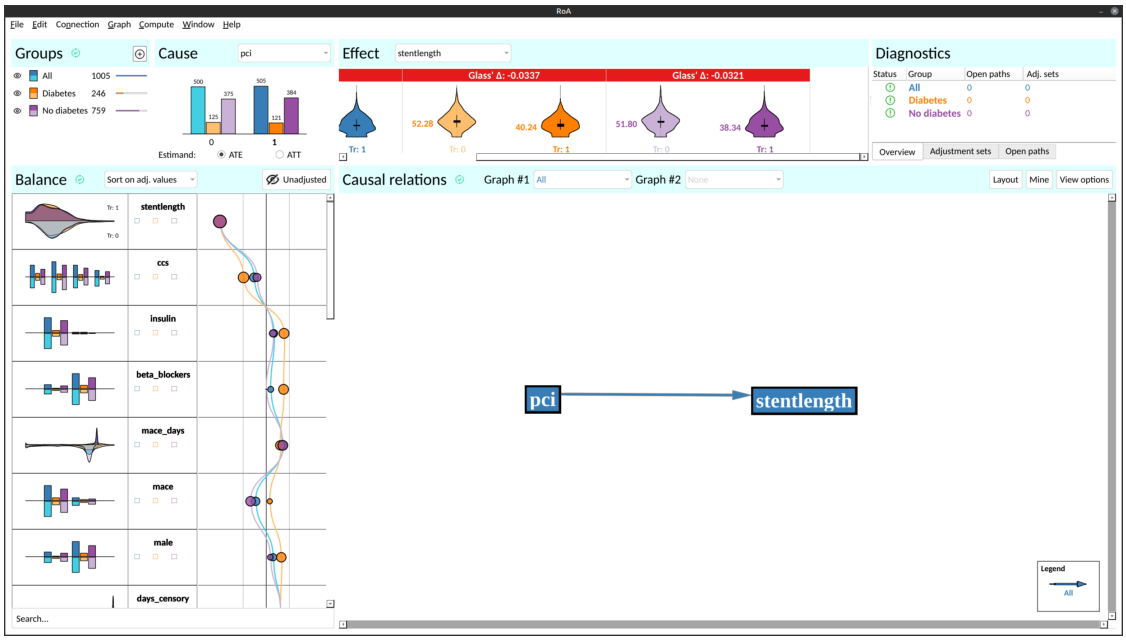


Figure 4: Subgroup analysis for the randomized treatment case. Results are shown for the subgroups in all panels using different colors for each group.

Observational study

If the treatment variable had not been randomized, the covariates would probably not be balanced. In that case, we could not have guaranteed an unconfounded treatment

effect estimation. Hence, we need to conduct an observational study instead of an RCT (Figure 2) to pursue unconfounded treatment effect estimation in this case. To illustrate this, we have purposely introduced selection bias in the study population, using conditional sampling, and used this adapted study population for our following scenarios. As a result, those patients with a higher risk for CAD and who underwent anatomy-guided PCI are more prevalent. More specifically, these patients exhibited at least two risk factors for CAD (hypercholesterolemia, diabetes, smoking, hypertension, or CAD as part of the family history).

The new situation is shown in Figure 5. Now we observe a highly skewed treatment variable distribution and a more considerable absolute difference in treatment effect. Additionally we notice multiple unbalanced variables: *hypertension*, *ace_inh*, *statins*, *diabetes*, *mace_days*, and *ca_anta*.

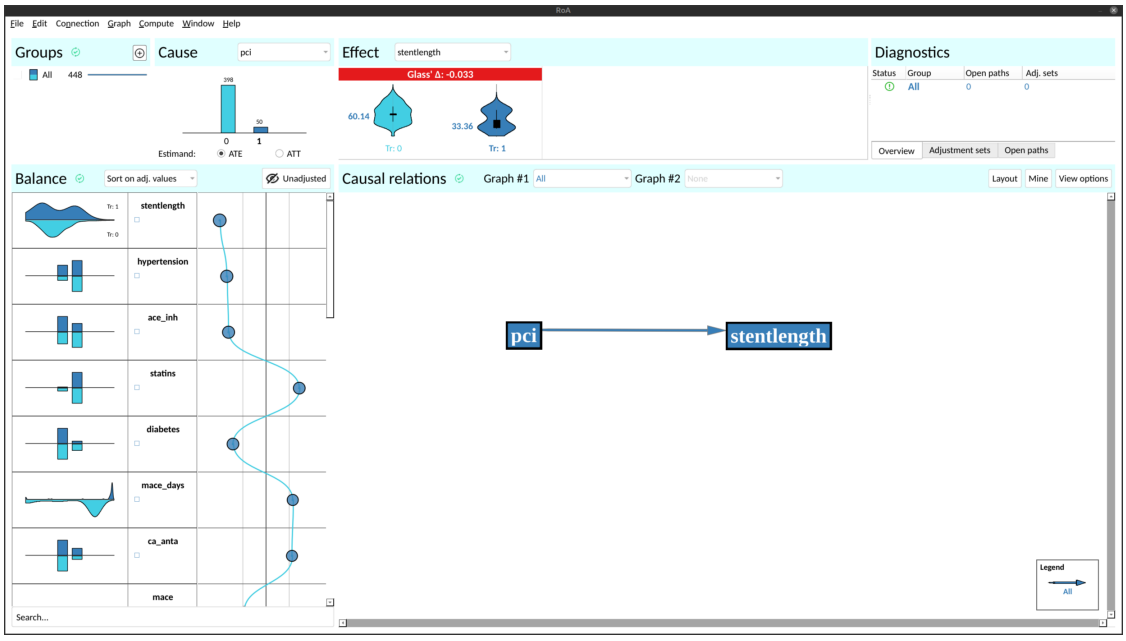


Figure 5: Biased dataset after conditional sampling. This time multiple covariates had SMD values outside of the acceptable range.

Manual adjustment

We decided to manually adjust all the previously observed unbalanced variables, except for mace days, which is considered to be unrelated by the experts. This scenario includes steps (a), (b), and (c), with an emphasis on step (b), shown in Figure 1. The result is shown in Figure 6. Now, the selected variables have been balanced, but this has caused the variables *nitrates*, *smoking*, and *css* to become unbalanced instead. In response, we added these three variables to obtain new results (see Figure 7). This time, all variables seem balanced and we accept the possibly confounded treatment effect estimate, which shows a statistically significant reduction of 29.76 mm in total average stent length (Glass's delta: -0.0648).

Adjustment based on graph analysis

Before we can analyze our causal graph, we have to define it. For a start, we employ

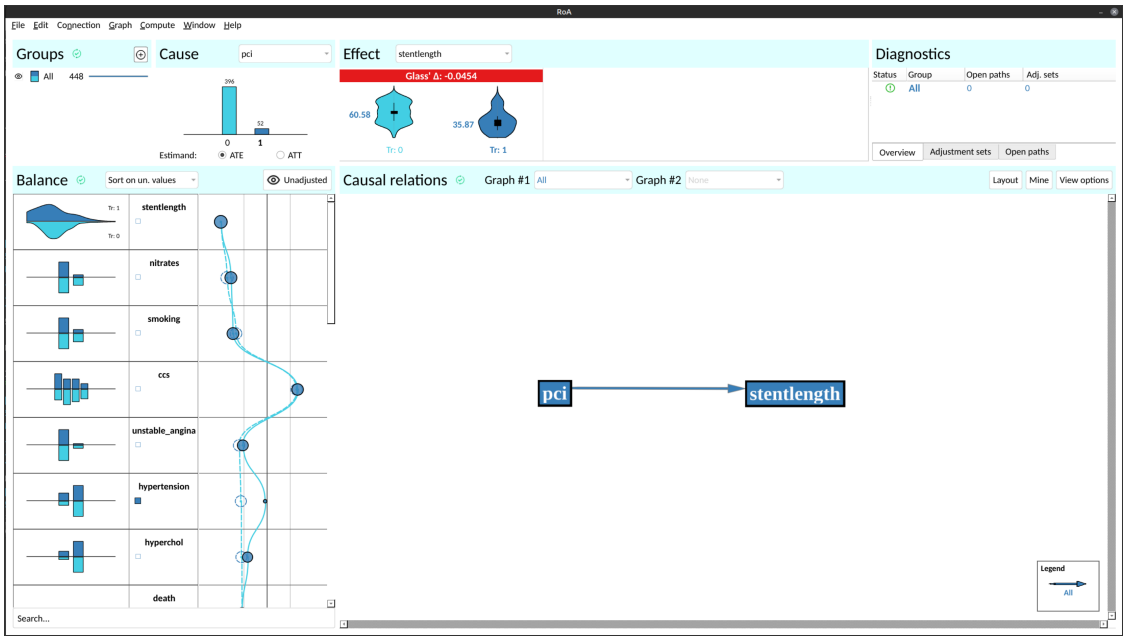


Figure 6: Biased dataset after conditional sampling with manual adjustment of covariates. The adjustment caused covariates outside of the adjustment to become unbalanced.

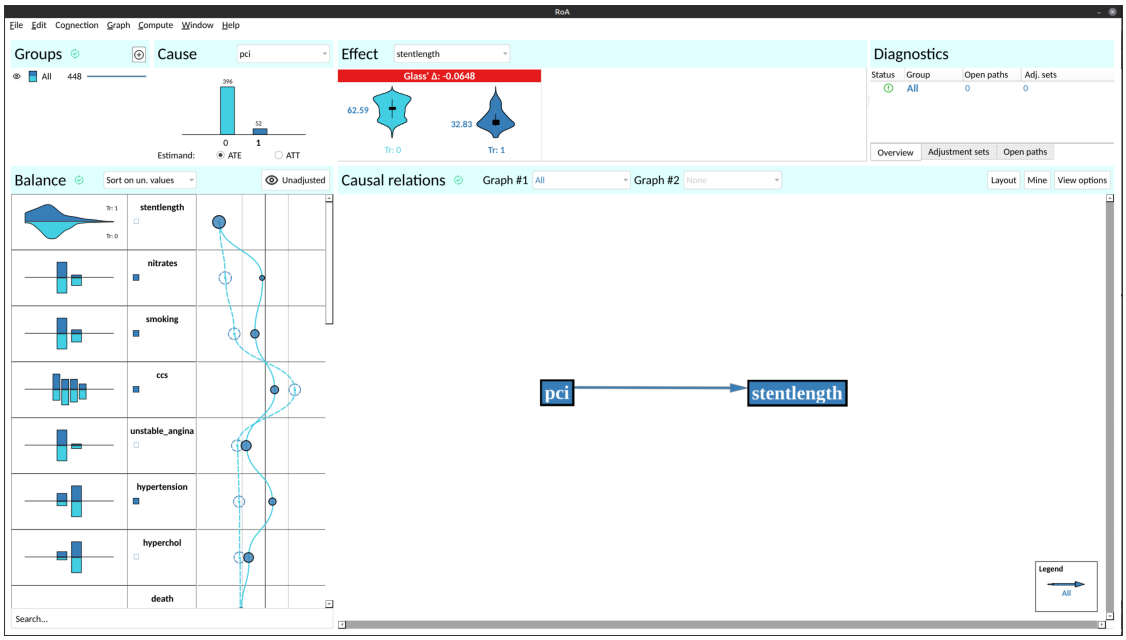


Figure 7: Biased dataset after conditional sampling with manual adjustment of covariates. The SMD values of all covariates are now within the acceptable range.

our mining algorithm on the *entire* study population with different cut-off values (of α) for significance testing of the arrows as shown in Figure 8 and Figure 9. This comprises step (d) shown in Figure 1. These mined graphs were shown to the experts, who defined the final expert graph based on all input and knowledge (step (e) and (f) in Figure 1). The difference visualization of the expert graph with the mined graph shown

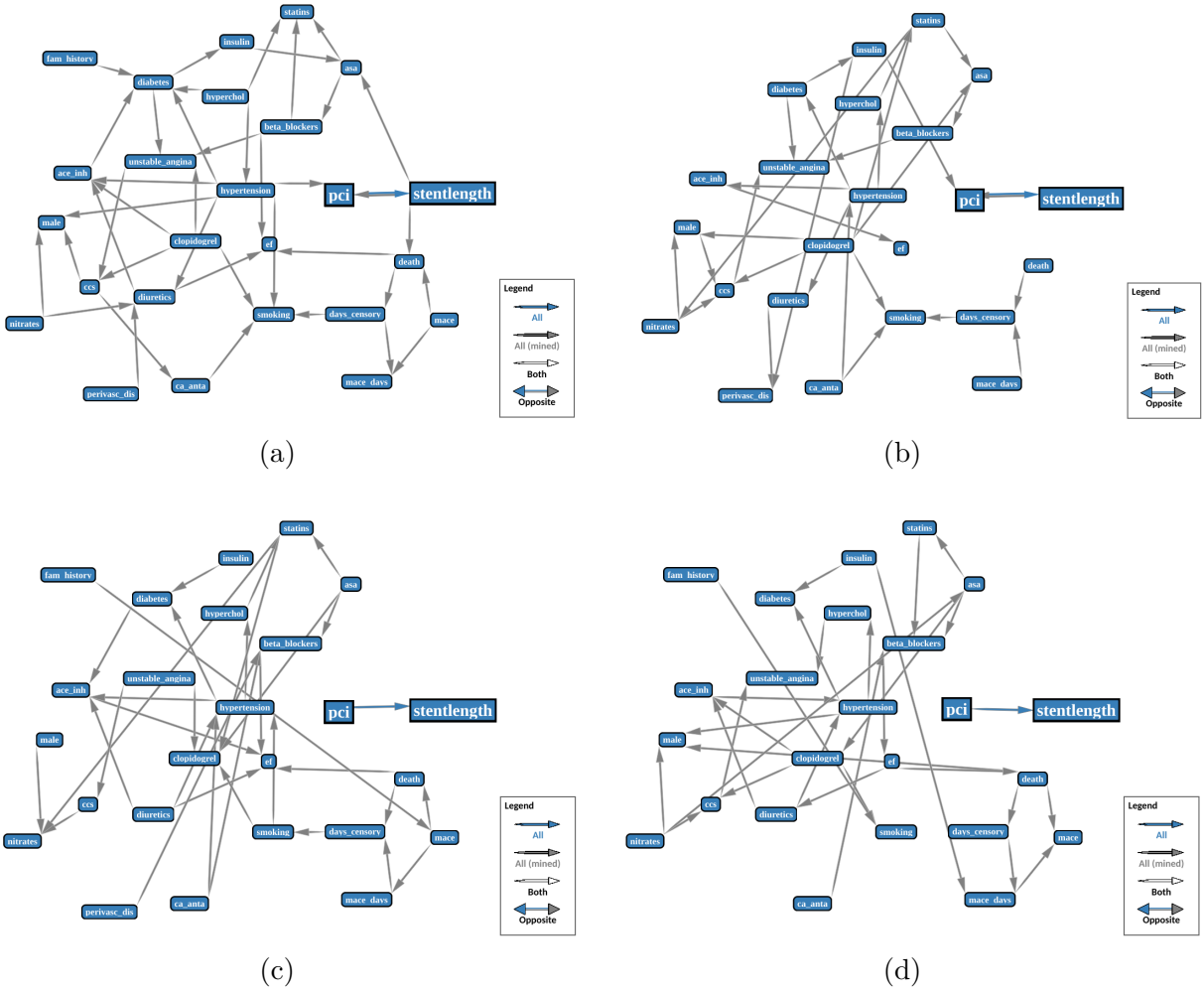


Figure 8: Several executions of the mining algorithm with two different α cut-off values for filtering arrows based on their significance. (a) Result using $\alpha \leq 0.1$ (b) First result for $\alpha \leq 0.05$. (c) Second result for $\alpha \leq 0.05$. (d) Third result for $\alpha \leq 0.05$.

in Figure 9d is shown in Figure 10.

Part of the reasoning followed for defining the expert graph is as follows. Cardiologists use x-ray images to decide to treat a patient to perform a PCI by looking at the diameter changes (narrowing) of the coronary arteries (anatomy-guided treatment), which is an inaccurate indicator of the lack of oxygen and nutrients. However, it is also possible to measure the pressure drop caused by a narrowing and objectively decide whether it causes limited blood flow (physiology-guided treatment). It is known that several covariates increase the risk of coronary artery disease. These risk factors include male sex, diabetes, smoking, family history of CAD, high blood pressure (hypertension), and high cholesterol in the patient's blood (hypercholesterolemia). Furthermore, medication is prescribed to reduce the progression of coronary artery disease. These include beta-blockers, calcium antagonists, ace inhibitors to lower blood pressure, statins to lower the level of cholesterol in the blood, and insulin for diabetes. Some drugs reduce the risk of myocardial infarction (aspirin and clopidogrel) and directly relieve chest pain (unstable angina, ccs) by using nitrates. The ejection fraction (ef) is a general

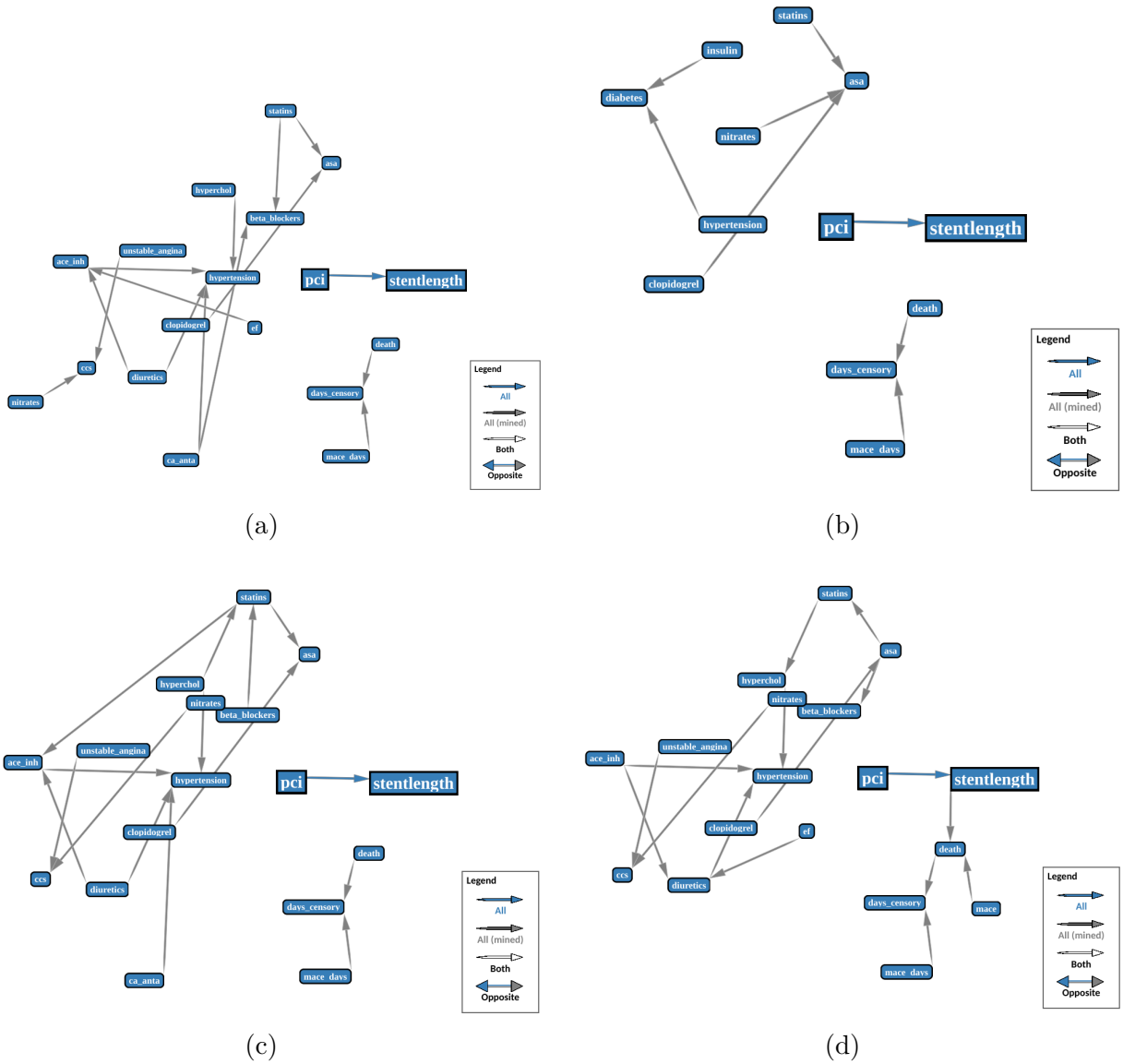


Figure 9: More executions of the mining algorithm with two stricter α cut-off values for filtering arrows based on their significance. (a) First result using $\alpha \leq 0.005$ (b) Second result for $\alpha \leq 0.005$. (c) First result for $\alpha \leq 0.001$. (d) Second result for $\alpha \leq 0.001$.

parameter to describe the patient's health status (and for predicting future cardiac events like a heart attack, stroke, or death). Patients with a low ejection fraction often suffer from heart failure for which a drug is prescribed to relieve heart failure symptoms like shortness of breath and fluid retention (diuretics).

All of these factors play a role in the extent of coronary artery disease and influence the type of PCI to perform. The influence on the treatment choice is removed by performing a randomized controlled trial like the FAME study. This is reflected in the balance panel shown in Figure 3. In addition, when mining graphs using various cut-offs for significance testing (α values) for filtering relations, there are relations that appear between *unstable_angina*, *ccs*, and *nitrates* that can be explained (see above). Also, relations between high blood pressure (*hypertension*) and a number of drugs

can be observed. The same holds for *diuretics*, *ef*, *insulin*, and *diabetes*. Note here that there are minimal to no arrows from any covariate pointing to PCI due to prior randomization. However, our expert graph (Figure 10) includes risk factors like *male*, *hyperchol*, *hypertension* and *diabetes*, which are confounding covariates that relate both to PCI and *stentlength*.

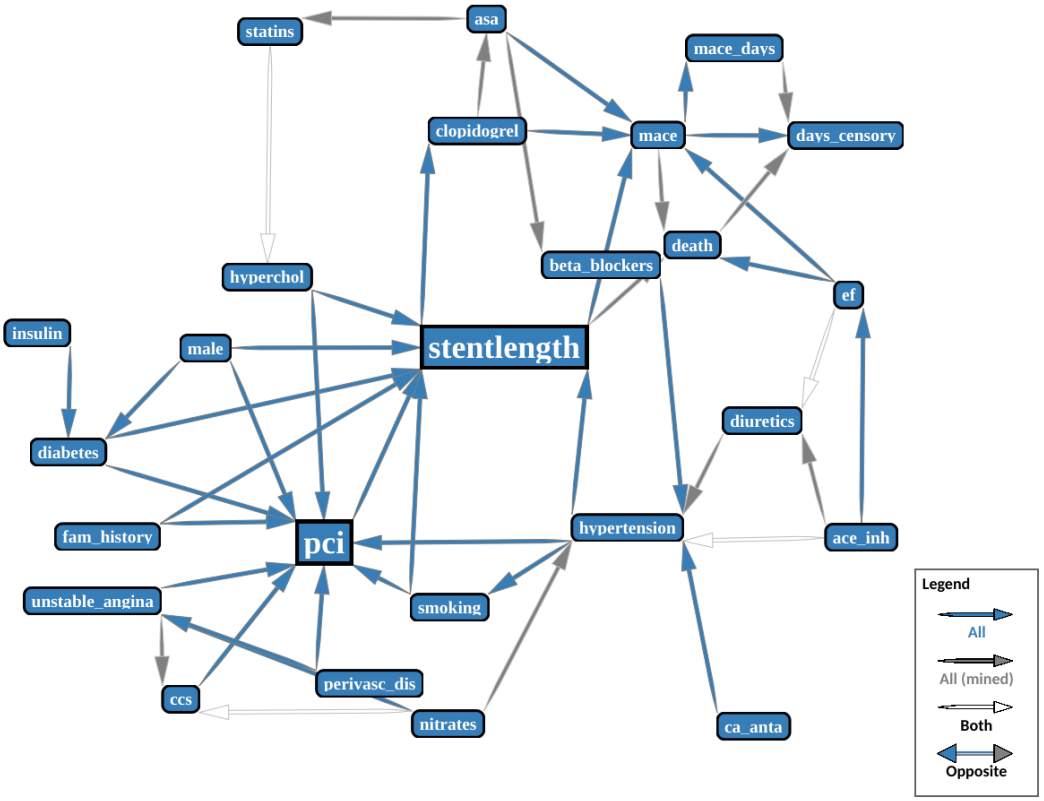


Figure 10: A difference visualization of the expert graph with the mined graph shown in Figure 9d.

With the causal expert graph defined, we proceeded with inspecting the diagnostics panel (see Figure 11). The panel indicated 15 open backdoor paths and one possible adjustment set: {*diabetes*, *fam_history*, *hyperchol*, *hypertension*, *male*, *smoking*}, which was selected (step (g) in Figure 1). The balance panel in Figure 11 presents both the adjusted and unadjusted values for comparison. Note that *ccs* is not part of the adjustment set and remains unbalanced. The new treatment effect estimation is a statistically significant reduction of 27.6 mm in total average stent length (Glass's delta: -0.055).

Subgroup analysis

In the previous scenario, we conveniently found an adjustment set to apply. Generally, however, this is not automatically the case. When no adjustment set is possible, performing subgroup analysis is another option. In this way, we can block the influence of a variable at the expense of obtaining multiple models designed for particular subgroups instead of the entire study population. During the expert discussion around the definition of the causal expert graph (step (e) and (f) in Figure 1), we encountered a version that did not yield any adjustment sets, shown in Figure 12a. The difference with our

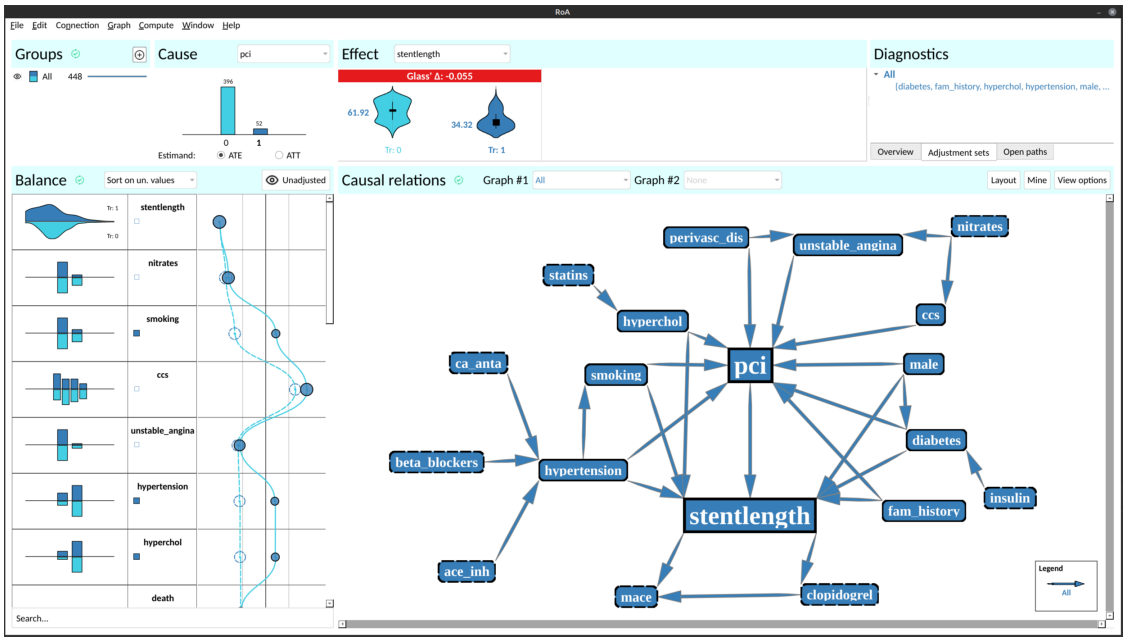


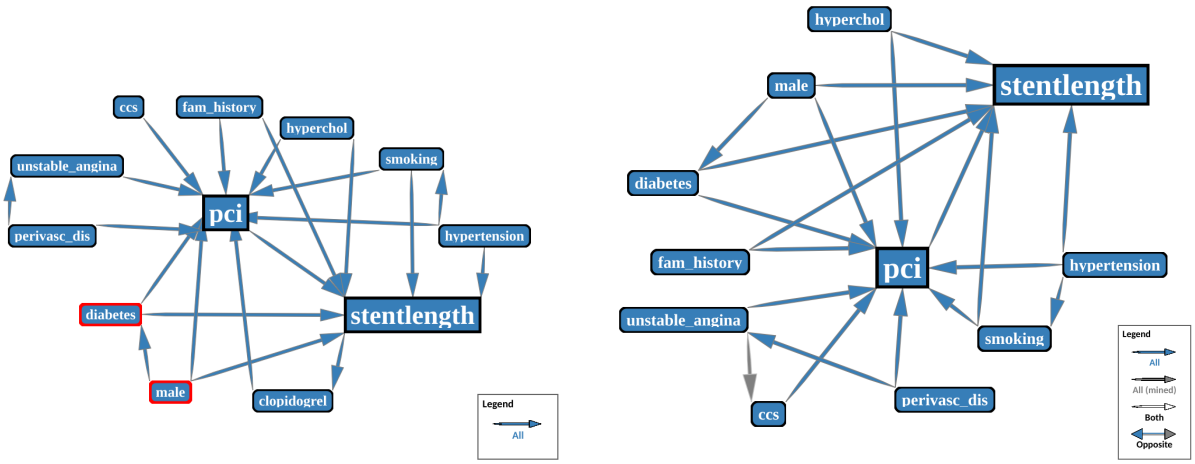
Figure 11: Application of computed adjustment to obtain a new (and deconfounded) treatment effect estimation. The causal graph shown was reduced to its Markov Blanket plus directly connected nodes.

final expert graph (Figure 12b) is just a single relation, namely *clopidogrel* $\rightarrow$ *PCI*. In the diagnostics panel, we observe that for the alternative causal expert graph, 19 open backdoor paths had been found with no possible adjustment set (see Figure 13a and Figure 13b).

In response, we created two subgroups based on the variable *clopidogrel* (step (h) in Figure 1), meaning one subgroup was administered the drug while the other was not. Furthermore, for both subgroups specifically, we removed the relationship *clopidogrel* $\rightarrow$ *PCI* from their causal graphs (step (d) in Figure 1) because now we assume no influence is being exerted anymore through that relationship. Consequently, the causal graphs for these subgroups are now made identical again to our expert graph for which the adjustment set is possible that we found earlier (see Figure 13c and Figure 13d). In this new situation, we have applied the adjustment sets for the subgroups (step (g) in Figure 1) to obtain separate treatment effect estimates for them, as shown in Figure 14.

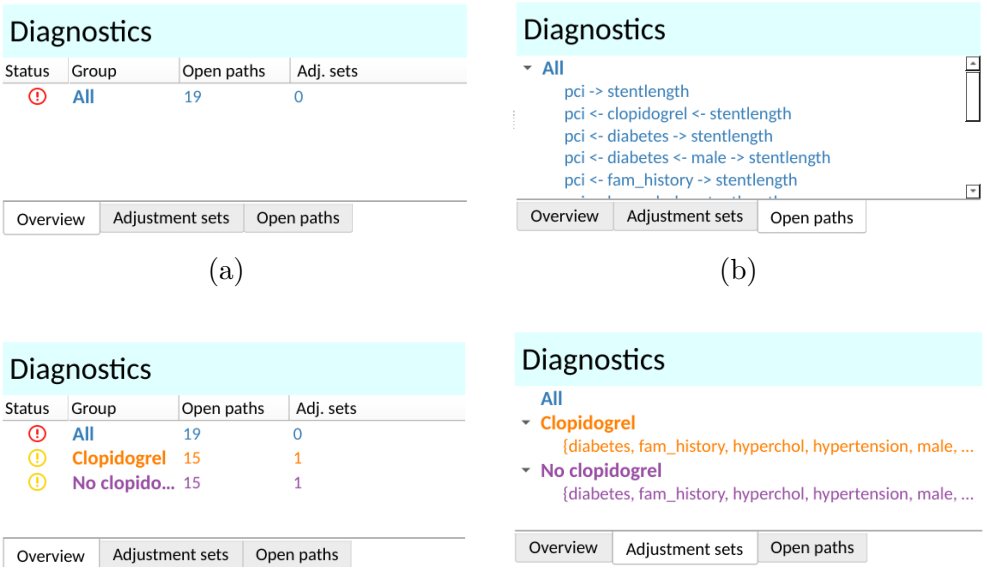
Estimates and insights

The final step shown in Figure 2 is obtaining the estimates and deriving insights. We have found that the treatment effect is independent of diabetes, based on the dataset used for the RCT. Because we intentionally introduced bias for the OBS to illustrate the scenarios supported by RoA, we cannot gather new clinical insights based on the analysis. On the other hand, it helped to show the connection between conducting an RCT and an OBS and the scenarios these entail.



(a) (b)

Figure 12: (a) The Markov Blanket of an alternative causal expert graph to which the relation *clopidogrel* $\rightarrow$ *PCI* has been added. This graphs does not allow for any adjustment sets. One of its 19 (open) backdoor paths is highlighted: *PCI* $\leftarrow$ *diabetes* $\leftarrow$ *male* $\rightarrow$ *stentlengh*. (b) The Markov Blanket of our expert causal graph (compared with the mined graph shown in Figure 9d).



(a) (b)

(c) (d)

Figure 13: Several situations indicated by the diagnostics panel. (a) No adjustment set is possible based on the alternative expert causal graph. (b) Some of the 19 open backdoor paths yielded by the alternative expert causal graph. (c) The subgroups based on *clopiogrel* both have 15 open backdoor paths and one possible adjustment set. (d) The only possible and identical adjustment set for the former subgroups: {*diabetes*, *fam_history*, *hyperchol*, *hypertension*, *male*, *smoking*}.

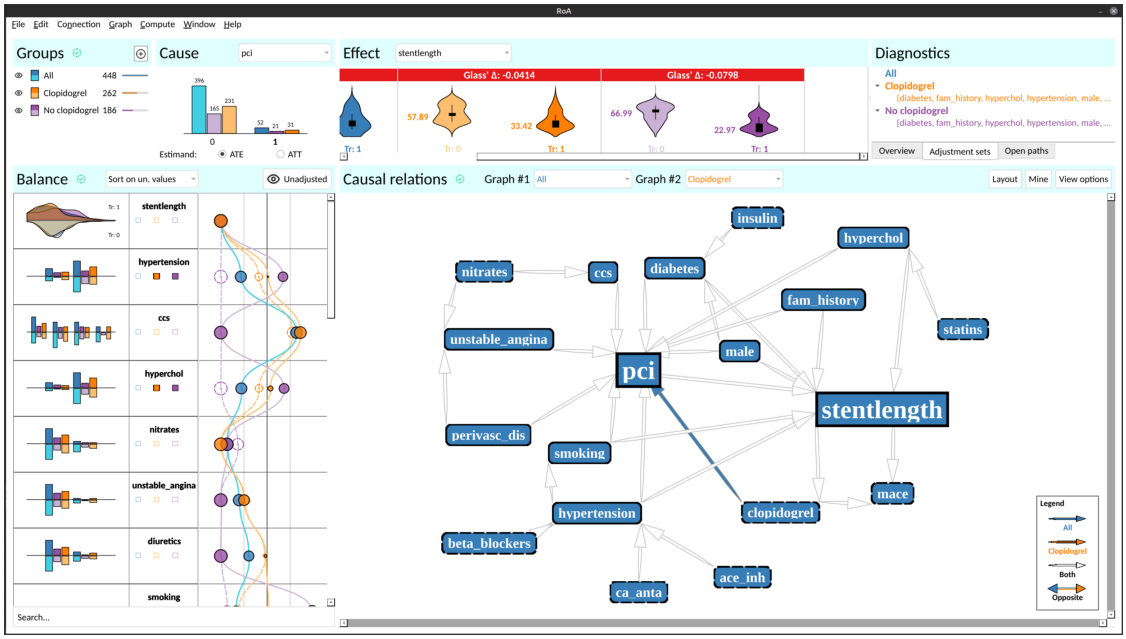


Figure 14: Subgroup analysis based on the variable *clopiogrel*. The adjustment sets for the subgroups have been applied. In the causal relations panel, the blue arrow highlights the (only) difference between the causal graph for the entire study population and the subgroup that was administered clopidogrel.