SUPPLEMENTAL MATERIAL

Data S1.

Two-Stage Instrumental Variable Methodology

In this work, we used the two-stage residual inclusion individual frailty (2SRI-F) algorithm.[1] We conducted two comparisons (PTA vs. atherectomy and stent vs. atherectomy) through two different IVs. The algorithm is as follows:

1. **Compute the IVs as the preference of using one treatment (T1 or T2) vs the reference (T0).** For measuring treatment preference in one moment, we considered the surgeries done in the same facility in the year (12 months) prior to a patients' procedure and computed the center-specific historical proportion of surgeries using one particular therapy (e.g. number of stent procedures / total number of stent and atherectomy procedures). We saved this proportion (denoted IV1 and IV2) and the number of surgeries performed in each hospital the last year, denoted VL.

2. **[First Stage]**. We performed a standard linear regression model to estimate the parameters of the treatment assignment model. In this model, we included the IV for the given treatment comparison and all measured covariates including the total number of surgeries performed to account for the relevant surgical experience of the hospital. That is,

$$T_{ji} = \hat{\alpha}_{j0} + \hat{\alpha}_{ji} IVj_i + \hat{\beta}_{j1} Z_{1i} + \hat{\beta}_{j2} Z_{2i} + \ldots + \hat{\beta}_{jK} Z_{Ki} + \hat{\gamma}_j VL_i \qquad [1]$$

where T_j is a binary random variable indicating which of the treatment j (j=1,2) and treatment 0, IVj is the instrumental variable relevant to these two treatments, $Z_1, \ldots, Z_K$

are K measured covariates and VL is the total number of relevant surgeries performed in the past 12-months.

3. We saved the residuals from the previous model: $R_{ji} = T_{ji}$-$\widehat{T_{ji}}$.
4. **[Second Stage]** We performed a proportional hazards Cox regression model with individual frailties including the covariates in [1] and the residuals, R_j.

We performed the 2-SRI-F procedure twice, once for the PTA vs. atherectomy comparison using IV1, and again for the stent vs. atherectomy comparison using IV2.

R code

1. Computing the IV for comparing treatments T1 and T2 [N is the sample size]

```
IVP1= sapply(1:N,function(i) {I=which(data$center==data$center[i] & as.numeric(data$start[i]-data$start)<= 365.24 & as.numeric(data$start[i]-data$start)> 0)  sum(data$trt[I]==1)/sum(data$trt[I]==1 | data$trt[I]==2)})

data$nprev1= IVP[1,]

data$iv1= IVP1
```

2. First Stage. [We adjusted data.t1 to just include two considered therapies and tr2 is defined appropriately]. Notice that the sample size, n1, just considers the two treatments.

```
S1= lm(trt2 ~ iv1 + race + age + ⋯ + htn + nprev1, data=data.t1)

data.t1$PRE= as.numeric(predict(S1,data.t1))

data.t1$RES= data.t1$trt2 - data.t1$PRE
```

3. Second Stage. [survival package is required]

```
tsA<-  coxph( Surv(timeAny,eventAny)~ trt2 +  race + age + … + RES + nprev1 +
frailty(1:n1,dis="gauss"), data=data.t1)
```

Data S2.

Instrumental Variable Assumption Assessment

The generalizability and validity of our IV findings depends on the strength with which we can make three key assumptions about our instrument. These assumptions are that our instrument: 1) has a causal effect on the exposure 2) only affects the outcome through the exposure 3) does not share common causes with the outcome. If these assumptions are held, then the effect we observed can be causal.[2] We found that our instruments were strongly associated with our exposure, treatment type, as evidenced by the large F-statistic values and increasing use of atherectomy for patients who receive treatment at centers with a high proportion of atherectomy procedures. The other IV assumptions cannot be verified from the data; hence, we relied on the expert knowledge of vascular surgery across our team to identify any potential assumption violations. Because our instrument is so strongly related to the exposure, any proposed alternative link between the instrument and outcome was ultimately related through the treatment type. We included total procedural volume as a covariate in our IV analyses to help justify the assumption that a hospital's experience with a given procedure is unrelated to patient outcomes after conditioning on observed covariates. Conditioning on total volume stops the presence of a general surgical volume learning effect from violating the third assumption, making a procedure specific learning effect the only threat to the validity of the IV. There is no evidence in the literature of a procedure-specific learning effect for endovascular PAD treatment and long-term outcomes. Thus, after careful consideration of each assumption, we are confident in the validity of our instrument and IV results.

Table S1. CPT codes used to identify outcomes in Medicare Claims

Outcome	CPT Codes				
Major Amputation	27590	27592	27881	27882	28805
	27591	27880			
Any Amputation	Major amputation codes +				
	28810	28820	28825		
Major Adverse Limb Event (major amputation OR reintervention)	Major amputation codes +				
	35521	35565	35305	35681	35456
	35351	35621	35306	35682	35459
	35355	35623	35371	35683	35470
	35361	35637	35372	35879	35474
	35363	35638	35533	35881	35483
	35537	35646	35556	35883	35485
	35538	35647	35558	35884	35495
	35539	35651	35566	35452	37205
	35540	35654	35571	35454	37206
	35541	35661	35583	35472	37208
	35546	35663	35585	35473	36200
	35539	35665	35587	35481	36245
	35548	35302	35656	35482	36246
	35549	35303	35666	35491	36247
	35551	35304	35671	35492	36248
	35563				

Supplemental References:

1. Martinez-Camblor P, Mackenzie T, Staiger DO, Goodney PP, O'Malley AJ. Adjusting for bias introduced by instrumental variable estimation in the Cox proportional hazards model. *Biostatistics (Oxford, England)*. 2019;20:80-96.
2. Lousdal ML. An introduction to instrumental variable assumptions, validation and estimation. *Emerging Themes in Epidemiology*. 2018;15:1.