The wicked path of causal inference in observational studies

Daniele Poole[1*] , Paul H. Mayo[2] and Morten Hylander Møller[3]

In this issue of Intensive Care Medicine [1], Arabi et al. present a pre-planned sub-study of the randomised clinical trial Pneumatic Compression for Preventing Venous Thromboembolism (PREVENT) [2]. In the PREVENT trial, use of adjunctive intermittent pneumatic compression combined with pharmacologic thromboprophylaxis with unfractionated heparin or low-molecular-weight heparin was compared to pharmacologic thromboprophylaxis in 2003 adult critically ill patients. In brief, PREVENT trial demonstrated that adjunctive intermittent pneumatic compression did not result in a lower incidence of proximal lower-limb deep-vein thrombosis (DVT) than pharmacologic thromboprophylaxis alone. In this reported pre-planned sub-study, the association between surveillance ultrasonography for DVT twice weekly and outcome was assessed. The surveillance group comprised 1682 patients enrolled in the PREVENT trial from 10 study sites, whereas the non-surveillance group consisted of 383 eligible non-enrolled patients. Outcomes in the two groups were compared using generalised linear mixed models and Cox proportional hazards models.

Surveillance ultrasonography was found to be associated with a statistically significant reduction in 90-day mortality in the Cox models, but not in the generalised linear mixed models (hazard ratio 0.75, 95%-confidence interval 0.57–0.99, p 0.04; odds ratio 0.87, 95%-confidence interval 0.63–1.20, p 0.39, respectively). Both models were consistent in indicating a statistically significant earlier diagnosis of DVT and pulmonary embolus (PE)

and an increase in the diagnosis of DVT, while no change in the frequency in diagnosis of PE was detected.

To demonstrate a causal relation between surveillance and mortality reduction a very robust statistical analysis was required, since we are dealing with explanatory purposes of the model [3]. We need, in this case, to compensate unbalances between the study arms in terms of prognostically important variables. To do this, we should include in the multivariable model all those variables that we know may be both prognostically important and unevenly distributed in the two groups (Fig. 1). The fact that the final model included only three prophylactic treatments, the presence of a femoral CVC at randomization, and the APACHE II score suggests that it was indeed too simplistic.

Further, when dealing with a treatment it is advisable to balance the probability of receiving it or not, with appropriate statistical tools such as propensity scores, which was not done in the study [4].

The analyses were even more complex for the outcomes of DVP and PE, since time of exposure to risk factors were not available in the control group, and could not be accounted for by the models. For example, the time a central line is in site is important in terms of DVT risk. Including in the model femoral CVC as a yes/no variable corresponds to comparing patients with a few days of exposure with those with prolonged exposure.
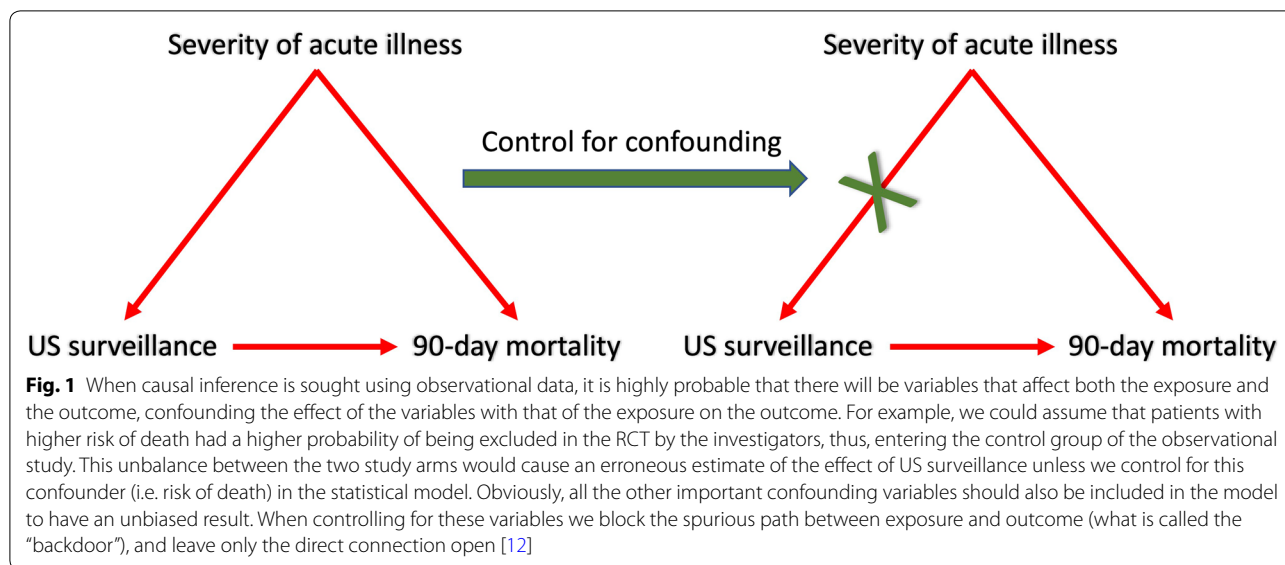
Interestingly, the 90-day Cox model indicated a protective effect of the intervention while the multilevel model did not, with the latter probably being more reliable since it accounted for centre variability, and hereby roughly controlling for case-mix and healthcare provision heterogeneity. The models, instead, were consistent in indicating an earlier diagnosis of DVT and PE.

The authors suggest in their conclusion that ultrasound surveillance may cause a 25% relative mortality reduction. Should we trust this finding? The insufficient number of

*Correspondence: daniele.poole@alice.it
[1] Operative Unit of Anesthesia and Intensive Care, S. Martino Hospital, Belluno, Italy
Full author information is available at the end of the article

Springer

Fig. 1 When causal inference is sought using observational data, it is highly probable that there will be variables that affect both the exposure and the outcome, confounding the effect of the variables with that of the exposure on the outcome. For example, we could assume that patients with higher risk of death had a higher probability of being excluded in the RCT by the investigators, thus, entering the control group of the observational study. This unbalance between the two study arms would cause an erroneous estimate of the effect of US surveillance unless we control for this confounder (i.e. risk of death) in the statistical model. Obviously, all the other important confounding variables should also be included in the model to have an unbiased result. When controlling for these variables we block the spurious path between exposure and outcome (what is called the "backdoor"), and leave only the direct connection open [12]

variables available in the dataset, the low number of variables included in the statistical models, and the inconsistent findings of the analysis, call for extreme caution and limit this result to the field of hypothesis generation. Indeed, screening use of diagnostic tests exposes to the risk of surveillance bias [5]; in our case, routine surveillance is more likely to identify DVT before it manifests clinically (if ever) but the prognostic implication of early diagnosis remains unclear.

Thus, although the results of the study are preliminary, intensivists may consider performing ultrasonography examination for DVT on their patients. Examination for DVT is within the scope of practice of intensivist as indicated in the International Expert Statement on Training in Critical Care Ultrasonography [6].

What is the feasibility and practical aspects of doing this? A DVT examination requires the intensivist to examine a series of well-defined anatomic levels of the veins of the leg from the popliteal to the upper common femoral artery level. Full compressibility of the target vein rules out thrombus at the site of the examination; visible thrombus or lack of compressibility are diagnostic of thrombus. Spectral and colour Doppler do not add diagnostic information to the compression study [7]. An adequate compression study of the relevant leg veins can be performed in a few minutes and can be readily incorporated into other aspects of the point of care ultrasonography examination [8]. Competence in performance of a DVT examination can be achieved with a short period of training [7]. The results of intensivist performed examinations are similar to those performed by specialists in radiology service, with the advantage of immediate results, cost efficiency, and ease of service [9].

In conclusion, we encourage ultrasound for early DVT diagnosis by intensivists to monitor critically ill patients, since it improves physicians' competence and provides a clearer picture of our patients. We should, however, be aware of the risks connected to screening procedures [10], which are not inherently harmful, but may trigger clinical choices not supported by evidence; in our case, the prescription of full anticoagulation for asymptomatic DVT.

To answer the question if surveillance ultrasonography improves outcome in adult critically ill patients, we need a high-quality RCT or, secondarily, an observational study based on an ad hoc prospectively designed database with a robust statistical analysis plan aimed at making causal inference [11].

Author details
[1] Operative Unit of Anesthesia and Intensive Care, S. Martino Hospital, Belluno, Italy. [2] Division of Pulmonary, Critical Care and Sleep Medicine, Northwell Health, Barbara and Donald Zucker School of Medicine At Hofstra/Northwell, Hempstead, NY, USA. [3] Department of Intensive Care Unit 4131, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark.

Compliance with ethical standards

Conflicts of interest
The authors declare they do not have any conflict of interest.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Arabi YM, Burns KEA, Alsolamy SJ, Alshahrani MS, Al-Hameed FM, Arshad Z, Almaani M, Hawa H, Mandourah Y, Almekhlafi GA, Aithan AA, Khalid I, Rifai J, Rasool G, Abdukahil SAI, Jose J, Afesh LY, the Saudi Critical Care Trials Group (2020) Surveillance or no surveillance ultrasonography for deep vein thrombosis and outcomes of critically ill patients: a pre-planned substudy of the PREVENT trial. Intensive Care Med. https://doi.org/10.1007/s00134-019-05899-1
2. Arabi YM, Al-Hameed F, Burns KEA, Mehta S, Alsolamy SJ, Alshahrani MS, Mandourah Y, Almekhlafi GA, Almaani M, Al Bshabshe A et al (2019) Adjunctive intermittent pneumatic compression for venous thromboprophylaxis. N Engl J Med 380:1305–1315
3. Katz MH (2003) Multivariable analysis: a primer for readers of medical research. Ann Intern Med 138:644–650
4. Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70:41–55
5. Sackett DL (1979) Bias in analytic research. J Chronic Dis 32:51–63
6. Expert Round Table on Ultrasound in ICU (2011) International expert statement on training standards for critical care ultrasonography. Intensive Care Med 37:1077–1083
7. Crisp JG, Lovato LM, Jang TB (2010) Compression ultrasonography of the lower extremity with portable vascular ultrasonography can accurately detect deep venous thrombosis in the emergency department. Ann Emerg Med 56:601–610
8. Blaivas M, Lambert MJ, Harwood RA, Wood JP, Konicki J (2000) Lower-extremity Doppler for deep venous thrombosis—can emergency physicians be accurate and fast? Acad Emerg Med 7:120–126
9. Kory PD, Pellecchia CM, Shiloh AL, Mayo PH, DiBello C, Koenig S (2011) Accuracy of ultrasonography performed by critical care physicians for the diagnosis of DVT. Chest 139:538–542
10. Zampieri FG, Einav S (2019) When will less monitoring and diagnostic testing benefit the patient more? Intensive Care Med 45:1447–1450
11. Poole D (2020) Power and challenges of big data: why clinical researchers should not be ignored. J Neurosurg Anesthesiol 32:3–5
12. Pearl J (1995) Causal diagrams for empirical research. Biometrika 82:669–710