



Universiteit
Leiden
The Netherlands

Personalization of medicine requires better observational evidence

Middelburg, R.A.; Arbous, M.S.; Middelburg, J.G.; Bom, J.G. van der

Citation

Middelburg, R. A., Arbous, M. S., Middelburg, J. G., & Bom, J. G. van der. (2018).
Personalization of medicine requires better observational evidence. *Clinical Epidemiology*, 10,
1391-1399. doi:10.2147/CLEP.S167137

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/86636>

Note: To cite this publication please use the final published version (if applicable).

Personalization of medicine requires better observational evidence

Rutger A Middelburg^{1,2}

M Sesmu Arbous^{2,3}

Judith G Middelburg⁴

Johanna G van der Bom^{1,2}

¹Center for Clinical Transfusion Research, Sanquin Research, Leiden, the Netherlands; ²Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands; ³Department of Intensive Care Medicine, Leiden University Medical Center, Leiden, the Netherlands; ⁴Department of Radiation Oncology, Erasmus University Medical Center, Rotterdam, the Netherlands

Abstract: Evidence-based medicine has become associated with a preference for randomized trials. Randomization is a powerful tool against both known and unknown confounding. However, due to cost-induced constraints in size, randomized trials are seldom able to provide the subgroup analyses needed to gain much insight into effect modification. To apply results to an individual patient, effect modification needs to be considered. Results from randomized trials are therefore often difficult to apply in daily clinical practice. Confounding by indication, which randomization aims to prevent, is caused by more severely ill patients being less or more likely to be treated. Therefore, the prognostic indicators that physicians use to make treatment decisions become confounders. However, these same prognostic indicators are also effect modifiers. This is in fact exactly why they are relevant to decision-making. We use simple, fictive numerical examples to illustrate these concepts. Then we argue that if we would record all relevant variables, it would simultaneously solve the problem of confounding by indication and allow quantification of effect modification. It has previously been argued that it is practically more feasible to “simply” randomize treatment allocation, than to adequately correct for confounding by indication. We will argue that, in the current age of evidence-based medicine and highly regulated randomized trials, this balance has shifted. We therefore call for better observational clinical research. However, careless acceptance of results from poorly performed observational research can lead clinicians seriously astray. Therefore, a more interactive approach toward the medical literature might be needed, where more room is made for scientific discussion and interpretation of results, instead of one-way reporting.

Keywords: treatment, personalized, effectiveness, effect modification, risk factors, confounding by indication

Introduction

Randomized trials are considered the gold standard for establishing effectiveness of treatments.¹ Therefore it is tempting to consider all other types of clinical research as “what you’re stuck with if a randomized trial is impossible.” To determine whether this, somewhat derogative, attitude toward observational research is justified, we should first consider why a randomized trial provides such highly reliable evidence. Proper consideration of the strengths of randomized trials will provide clear guidance on when a randomized trial is absolutely indicated. Conversely, it will also hint at when a randomized trial could be either unnecessary or even contraindicated.

It has previously been reasoned that confounding by indication, which randomized trials aim to prevent, could also be corrected for in data analyses if treatment

Correspondence: Rutger A Middelburg
Center for Clinical Transfusion Research,
Sanquin Research, Leiden, Plesmanlaan
1a, 2333 BZ, Leiden, the Netherlands
Tel +31 071 568 5060
Fax +31 071 568 5191
Email R.A.Middelburg@lumc.nl

allocation decisions could be adequately modeled.^{2,3} This was, however, considered very difficult. Therefore it was considered more practically feasible to avoid confounding by indication all together, by performing a randomized trial. Furthermore, it was proposed that allocation of treatment is likely to be subject to some latent processes (ie, unrecognized or unmeasurable: “gut-feeling”). It was suggested that confounding by indication, therefore, cannot be measured directly and fully but only tangentially, if it is recognized at all.⁴

However, since then, 40 years have passed and four major developments have shifted this balance dramatically. First, two and a half decades after the introduction of evidence-based medicine,^{5,6} physicians find themselves obligated to explore the rational arguments that must motivate each treatment decision. Therefore, the role of the immeasurable “gut-feeling” should be diminished completely. Second, increasing diagnostic testing and digitalization of test results has made the information physicians use for treatment decisions more readily available to researchers. Third, increasing regulation of randomized trials including monitoring, auditing, and traceability demands has greatly increased the logistical challenges of randomized trials. With these logistical challenges, the cost also increases and consequently, the size of trials must be minimized to maintain economic viability. Fourth, it has been increasingly recognized that individualization of the evidence base for treatment decisions is needed.⁷

The difficulties in applying results from randomized trials to individual patients have been described in some detail.^{8–10} This problem is largely related to the limitations in size of trials, as it could be solved by sufficiently subgrouping patients in the analyses. The potential impact of subgroup differences has been quantified by comparing the risk difference between high- and low-risk patients in 32 large randomized trials.¹¹ In these analyses, all but one showed a significant difference in treatment effectiveness between the highest and lowest risk quartiles.¹¹

Here we present fictive numerical examples concerning blood transfusion in the ICU to illustrate conceptual considerations instrumental to understanding both the power and the limitations of randomized trials. We argue that, in the current setting, randomization often comes at the price of a loss of personalization of medicine. In other words, randomization can be a powerful tool to help prove that a treatment can provide benefit to some, but it may also be what’s keeping us from proving in whom it will do so.

Why we randomize: confounding by indication

Every field of clinical medicine has some well-known examples of published observational research showing associations between treatment and outcome, which can be blamed on baseline incomparability between the treated and untreated groups.¹² It is the problem of comparing apples and oranges. Patients with worse prognosis are either more or less likely to be treated and the observed outcome is in spite of treatment, rather than due to treatment. This is the problem of confounding by indication.^{12,13} Confounding by indication is the reason we randomize.^{1,14–16}

Figure 1 illustrates the assumed true causal relation between blood transfusion and mortality in a fictive ICU population. In Figure 1, the entire population is depicted in two mutually exclusive situations, both nontransfused (left panel) and transfused (right panel). This is called a counterfactual representation of causality, because it is always “counter to the facts” (ie, it is impossible to observe in real life).^{17–19} Counterfactual comparisons are considered theoretically ideal for establishing causality, because they compare every individual in the treated situation to himself or herself in the untreated situation.^{17–19} Although even counterfactual theory has its limitations,²⁰ it roughly corresponds to our intuitive understanding of causality (ie, would the same outcome have occurred with treatment as without or vice versa).

For the simplified example in Figure 1, the population is divided into three different, equal-sized groups. The risk of death differs between these groups. If each group contains 200 patients and the entire population was left nontransfused, we assume the following risks of death for the three groups considered: 20% for patients with cardiovascular disease, 40% for patients with active bleeding, and 80% for patients with both cardiovascular disease and active bleeding. The true causal effect of blood transfusion is set at an absolute risk reduction of 10% for all patients for each of the three groups. This also corresponds to an average risk reduction of 10% for the entire population, in spite of a difference in baseline risk of death between the groups. This baseline difference in itself is insufficient to cause confounding by indication.

Confounding by indication arises from the combination of this difference in prognosis and a different treatment prevalence. Figure 2 illustrates both how confounding by indication can distort the true causal relation and how a randomized trial can estimate this relation correctly.

In panel A of Figure 2, the choice to treat or not to treat is left to the physician and the physician is more likely to trans-

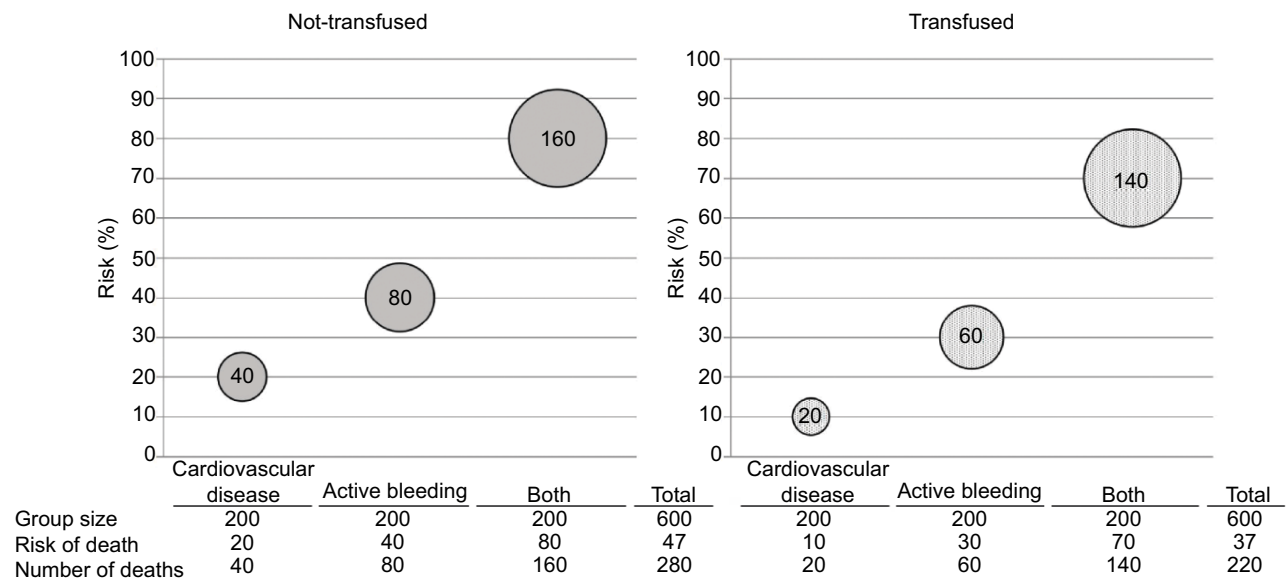


Figure 1 True causal relation between blood transfusion and mortality in a fictive ICU population, which is either completely nontransfused or completely transfused.

Notes: Schematic representation of the hypothetical effects of blood transfusion in patients with different risk factor profiles. Three different profiles are considered: the presence of cardiovascular disease, active bleeding, or both. Solid gray bubbles represent nontransfused patients and patterned bubbles represent transfused patients. Height of the position of the bubbles indicates the risk of death (%) and the size of the bubble represents the actual number of deaths. The table under each graph shows the represented numbers. The average risk of death (ie, the risk for the total population) can be calculated by dividing the total number of deaths (ie, added from all three subgroups) by the total population size. The risk of death differs between the groups. The effect of blood transfusion is a risk reduction of 10% for each of the three groups. As can be seen from the table under the graphs, this also corresponds to an average risk reduction of 10% (ie, from 47% to 37%). This average risk difference is not affected by the difference in baseline risk of death between the groups.

fuse patients with worse prognosis: 50 of 200 patients (25%) with cardiovascular disease, 100 of 200 (50%) with active bleeding, and 180 of 200 (90%) with both are transfused. These groups still also experience the previously assumed death risks of 20%, 40%, and 80% in the nontransfused patients. Although the absolute risk reduction for mortality is still assumed to be 10% for any of the three groups, the crude average risk difference for the total population is now an increase of 17%. This reversal of the true beneficial effect of blood transfusions into an increased risk of death occurs because the deaths from the higher risk groups (ie, active bleeding or both cardiovascular disease and active bleeding) weigh more heavily among the transfused patients, since the prevalence of transfusions is higher in these groups. In other words, the transfused group contains more seriously ill patients and is therefore not exchangeable with the nontransfused group.^{17–19} The table under the graph also shows how the true effect can be correctly estimated through statistical correction, which takes a weighted average of the stratum-specific risk reductions (ie, each stratum-specific absolute risk reduction is 10%, making the average also the true causal risk reduction of 10%, irrespective of the relative weights of the different strata). This approach creates what in counterfactual theory is called *conditional exchangeability*.^{17–19}

Conditional on having a certain disease, the transfused and nontransfused patients are considered *exchangeable*, in terms of prognosis. Put differently, there is no confounding within strata of the three measured confounders.

In this simplified example, correction for the prognosis of the patient is straightforward. Conversely, in real life, confounding by indication is notoriously difficult to adequately correct for.¹³ Many different factors determine a patient's prognosis and if any of these also influence a physician's decision to prescribe a certain treatment, these will cause confounding by indication. If we do not know exactly which factors a physician weighted in his or her decision to treat a patient, we cannot measure these factors. If we have not measured these factors, it will be impossible to apply a statistical correction.

Panel B of Figure 2 shows how randomization can fix the problem of confounding by indication, without the need for statistical correction. As mentioned above, confounding by indication only occurs if the prevalence of treatment differs between groups of patients with different prognosis. By randomly assigning patients to be either treated or nontreated, the probability of being treated is made independent of prognosis. Therefore, the prevalence of treatment is made equal between the groups of different prognosis. In the example

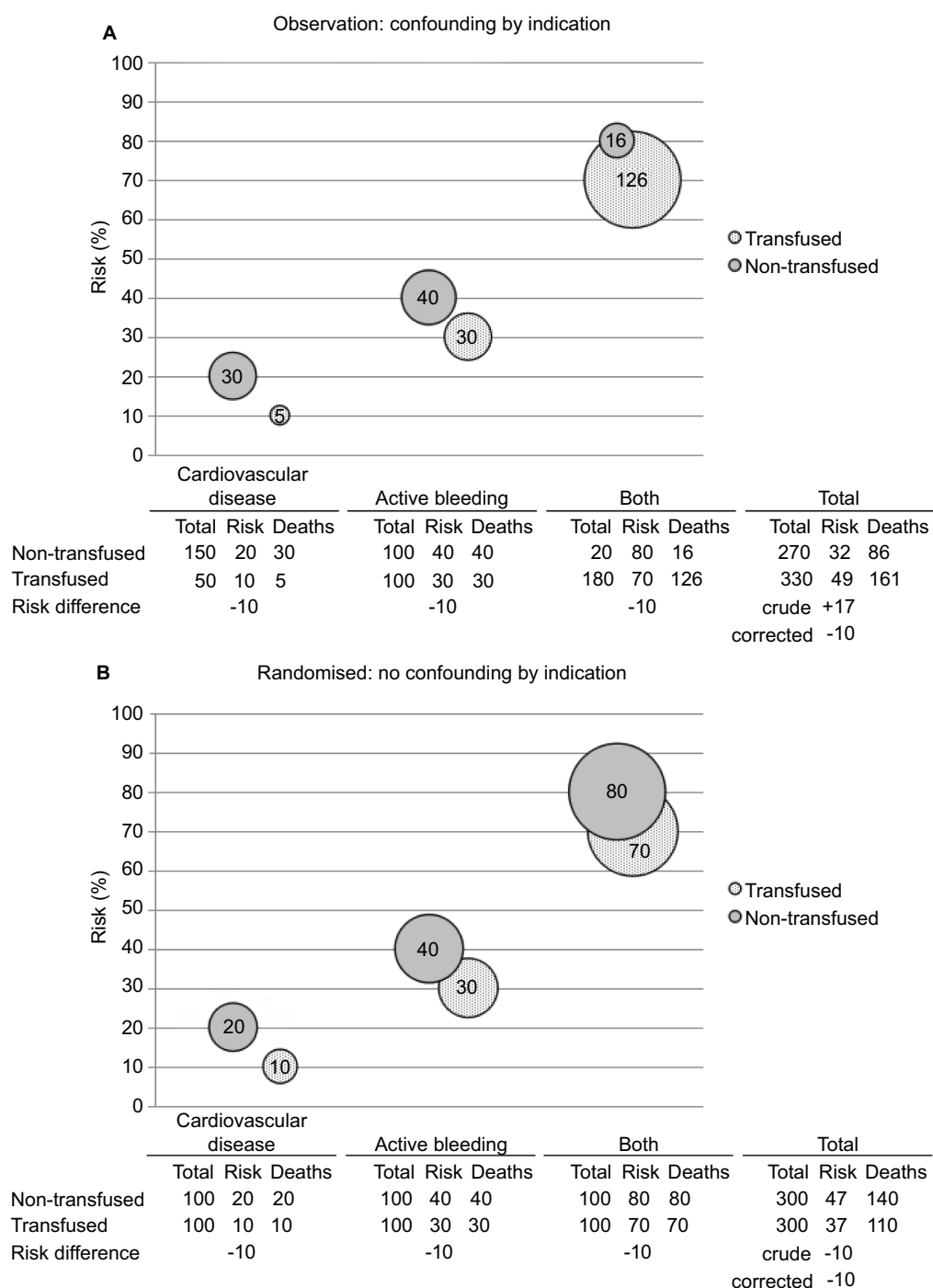


Figure 2 Causal relationship distorted by confounding by indication and correctly estimated in a randomized trial.

Notes: Schematic representation of the hypothetical effects of blood transfusion in patients with different risk factor profiles. Three different profiles are considered: the presence of cardiovascular disease, active bleeding, or both. Solid gray bubbles represent nontransfused patients and patterned bubbles represent transfused patients. Height of the position of the bubbles indicates the risk of death (%) and size of the bubble represents the actual number of deaths. The table under each graph shows the represented numbers. In panel **A**, in an observational study, there is confounding by indication. The true effect of transfusion (ie, a risk reduction of 10%) can still be estimated by statistical correction, but the crude analysis shows a risk increase of 17%. In real life, full statistical correction for confounding by indication is often impossible. In panel **B**, confounding is removed by randomization. Therefore, the crude analysis correctly estimates the true effect of blood transfusion.

they each contain 100 transfused and 100 nontransfused patients. This corresponds almost perfectly to the two situations depicted in Figure 1, except that all groups are half the size. We had to divide our population into two and transfuse only one half, because in real life, we cannot study the entire population in two mutually exclusive situations (ie, both nontransfused and transfused, such as depicted in Figure 1). Randomization is the closest we can possibly get to a real counterfactual situation.

The above example illustrates the reason we randomize. We do so to avoid confounding in situations where we cannot correct for it.^{1,14,15} Often, we attempt to correct for confounding by indication but are unsure if we fully succeeded in doing so. In these situations, randomization will, on average, over many studies, provide valid evidence of the effectiveness of a treatment without the danger of incompletely corrected confounding by indication.

A problem randomization just can't fix: effect modification

In spite of the tremendous power of randomized trials to provide us with highly reliable evidence of treatment effectiveness, there remains a problem that randomization just cannot fix. In all of the previous discussions, we have assumed the effect of blood transfusions to be the same for all three patient groups considered. This is of course highly improbable.²¹ In fact, an expected difference in effectiveness is precisely what causes confounding by indication. A physician is more likely to transfuse a patient whom he or she thinks is more likely to benefit from that transfusion. In other words, if 90% of patients with both cardiovascular disease and active bleeding are transfused and only 25% of patients with cardiovascular disease alone, the physician apparently expects that patients with both cardiovascular disease and active bleeding are much more likely to benefit from receiving a transfusion. This expectation is based on something—if not on an actual well-documented, systematically gathered and identifiable, and objectively verifiable evidence base, then at least on extensive clinical experience.²²

Whatever the exact basis for the clinical decision to transfuse one patient and not the other is, it is likely that the effect of transfusion will really be different between two patients. The effect will then be modified by the patient's risk profile.²¹ Although the presence of effect modification does not necessarily detract from the validity of the conclusion of a randomized trial, it is important to be aware of how it can influence the results.²³ To appreciate how effect modification

influences the results of a randomized trial, we continue our simplified example with the three patient groups in Figure 3.

Panel A of Figure 3 shows a more realistic outcome of the randomized trial also depicted in panel B of Figure 2 (ie, an outcome showing effect modification across subgroups of patients). In this example transfusion has no effect on mortality in patients with cardiovascular disease alone. In patients with active bleeding, transfusion reduces the absolute risk of death by 10% and in patients with both cardiovascular disease and active bleeding, the absolute risk of death is reduced by 30%. Since this is a randomized trial, we assume no confounding. Therefore, both the crude and corrected effects give valid estimates of the true effect in this population. This true effect is a population average reduction in risk of death of 13%.

Panel B of Figure 3 shows the target population to which the results of the randomized trial in panel A are going to be applied. As can be seen from the table under the graph, the absolute risk differences for each of the three groups are the same as observed in the randomized trial, reaffirming the validity of our estimates from that trial. For clarity we have assumed no confounding by indication in this target population. This can also be checked in the table under the graph where it is shown that all the three groups have 50% of their patients transfused. Therefore, prognosis and treatment prevalence are again independent, and there can be no confounding by indication.

There is only one difference between the target population and the randomized trial in Figure 3. The risk factor distribution is different. There are more patients with cardiovascular disease alone and fewer with both cardiovascular disease and active bleeding in the target population. Therefore, there are more patients experiencing no effect at all and fewer patients experiencing the maximum effect of 30% reduction in the risk of death. As a consequence, the population average effect is much smaller. This effect is a 5% instead of a 13% reduction in the absolute risk of death. This is probably one of the most important reasons why treatments applied in daily clinical practice can have disappointingly little effect compared to expectations raised in randomized trials.

It also illustrates what important information is typically missing from the interpretation of results from randomized trials. Subgroup analyses would have told us to transfuse only patients with active bleeding, whether alone or in combination with cardiovascular disease. However, subgroup analyses are generally discouraged on the basis of a lack of power and the presumed danger of post hoc analyses. To allow subgroup

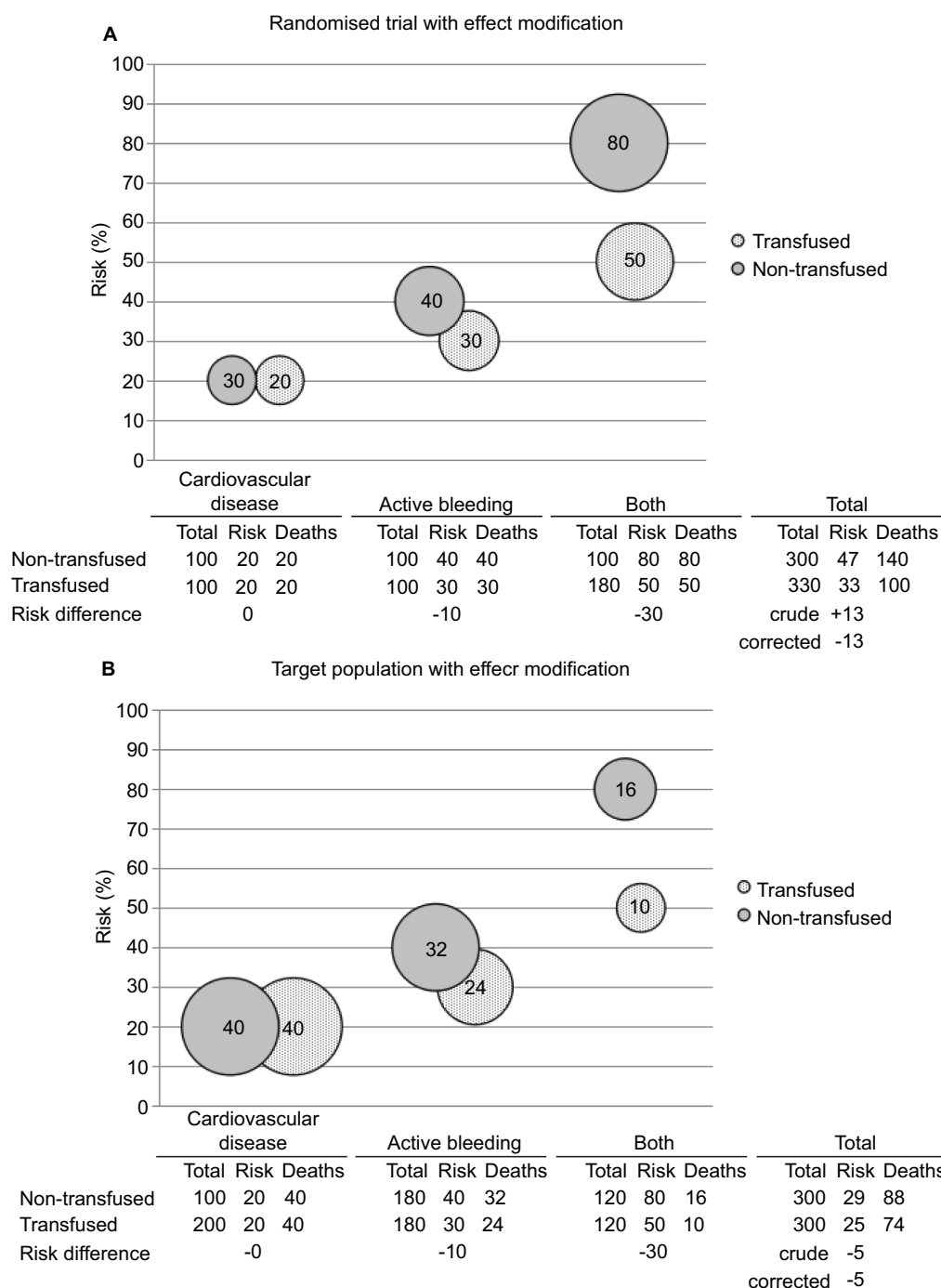


Figure 3 In the presence of effect modification, a randomized trial provides a completely valid effect estimate, which is, nonetheless, irrelevant for a target population with different risk factor distribution.

Notes: Schematic representation of the hypothetical effects of blood transfusion in patients with different risk factor profiles. Three different profiles are considered: the presence of cardiovascular disease, active bleeding, or both. Solid gray bubbles represent nontransfused patients and patterned bubbles represent transfused patients. Height of the position of the bubbles indicates the risk of death (%) and size of the bubble represents the actual number of deaths. The table under each graph shows the represented numbers. There is effect modification by subgroup. There is no confounding by indication (ie, the crude and corrected risk differences are the same). Still it is clear that the distribution of patients over the subgroups is crucial. In panel **A**, a randomized trial was performed and the estimated risk difference was -13%. Panel **B** shows the target population (still without confounding by indication), the real risk difference in that population would be -5%. However, these are both estimates of the average risk difference. What we really want to know is that transfusing patients with cardiovascular disease alone would be useless, while transfusing patients with both cardiovascular disease and active bleeding would result in a risk difference of -30%. The only way to know this is to perform subgroup analyses.

analyses, they have to be prespecified, which in turn means that a trial has to be sufficiently powered. This power should allow for three considerations. First, each subgroup should be of sufficient size, meaning the required trial size is inflated by a factor roughly equal, but usually bigger than the number of intended subgroup categories. Second, the power to show a difference in effect size between groups is usually smaller than the power to show the presence of an effect within a group. This even holds true if no effect is expected in one of the groups, because this lack of effect is estimated with a margin of uncertainty, while normal hypothesis testing takes the null hypothesis as a fact without uncertainty. This means a further inflation of trial size. Third, some kind of allowance will have to be made for multiple testing. All in all, the correct implementation of subgroup analyses in a randomized trial will cause the trial to be at least several times and often orders of magnitude larger than a similar trial without subgroup analyses. The way randomized trials are currently performed, including all regulations that have to be complied with, usually makes a trial that is sufficiently powered for extensive subgroup analyses financially impossible. Subgroups, in conclusion, are things better not meddled

with in randomized trials. Subgroups, unfortunately, are also exactly what we need to allow clinically meaningful interpretation of trial results.

If randomized trials are poorly suited for addressing effect modification and clinical decision-making needs reliable information on effect modification, can observational studies provide the necessary evidence? This question brings us back to our initial problem: confounding by indication.

Real-life situation: confounding by indication and effect modification

In real life, we are stuck with a seemingly intractable combination of two problems. First, we need an estimate free of the distortive effect of confounding by indication. This is the reason we do randomized trials. Second, we need to estimate subgroup-specific effects. For reasons of practicality, this is effectively impossible in randomized trials. We therefore seem to have arrived at an impasse.

However, the fact that the only reason to randomize is to avoid confounding in no way implies confounding by indication to be solvable only by randomization. As can be seen from Figure 4, there is no confounding by indication

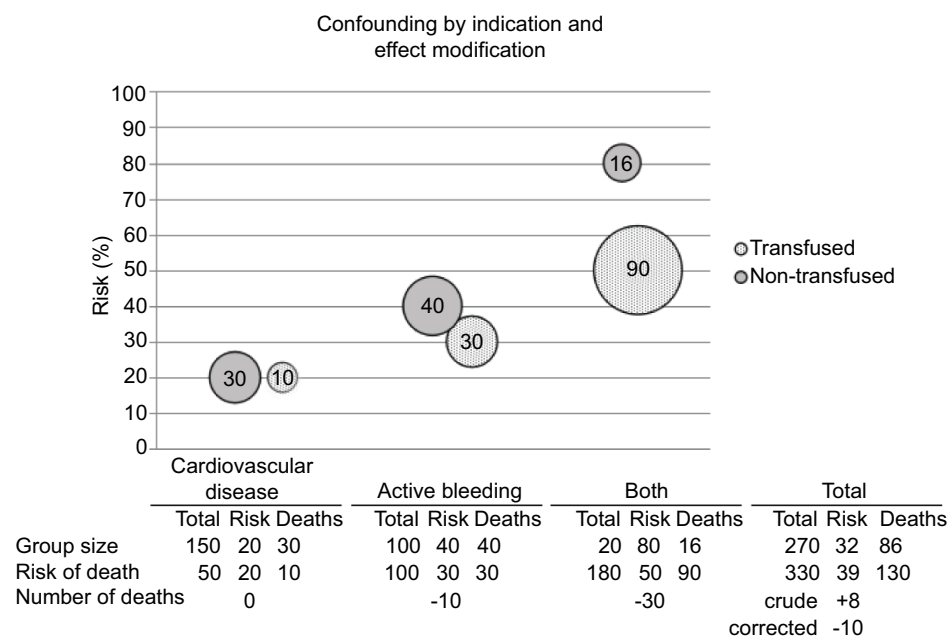


Figure 4 The most likely real-life situation: both confounding by indication and effect modification influence the results.

Notes: Schematic representation of the hypothetical effects of blood transfusion in patients with different risk factor profiles. Three different profiles are considered: the presence of cardiovascular disease, active bleeding, or both. Solid gray bubbles represent nontransfused patients and patterned bubbles represent transfused patients. Height of the position of the bubbles indicates the risk of death (%) and size of the bubble represents the actual number of deaths. The table under each graph shows the represented numbers. There are both confounding by indication and effect modification. This is the most likely scenario for most clinical situations. It is clear (from the relative positions of the bubbles and the risk differences in the table) that within strata of subgroup, this situation is comparable to the situation in Figure 3. To estimate the influence of effect modification, we need to perform subgroup analyses. If we perform subgroup analyses, confounding by indication becomes irrelevant, because within subgroup analyses risk difference is unaffected by this confounding (see also panel A of Figure 2).

within subgroups. Obviously, in this dramatically simplified example, it is easy to identify all three relevant subgroups, while in real life, there may be dozens of unknown confounders. However, as argued before, all these confounders are confounders because, and only because, the physician expects them to be effect modifiers. The physician is more likely to treat a patient in which he or she expects the biggest treatment benefit. In this way she turns prognostic indicators, which are associated with expected benefit (ie, expected effect modifiers) into real confounders. A critical physician should then demand evidence of this effect modification, since this is instrumental to informed clinical decision-making. It would be impossible to practice any form of evidence-based medicine without asking oneself “why should I treat this particular patient while I do not treat that other one?” This means effect modification has to be quantified and to do so, these confounders have to be identified, because we need subgroup analyses for each and every one of them. Once these confounders have been identified, such subgroup analyses are much more efficiently carried out in observational than in randomized trials. Simply because observational research can be carried out at much lower financial and ethical costs. Further, since these subgroup analyses will also solve the problem of confounding by indication, randomization is no longer needed anyway. Finally, identification of confounders relies entirely on physicians making explicit why they chose to treat one patient and not the other.

It may seem like a daunting task to identify all patient characteristics weighted in a treatment decision, but in this age of evidence-based medicine it is the only way to personalize our treatments. Without personalization, almost all patients will be overtreated or undertreated.²² It may therefore well be time to finally start seriously working at this daunting task, as it is the only way to truly personalize medicine in an evidence-based way.

Discussion

In the above, we have argued that, although randomization is an immensely powerful tool against confounding by indication, personalization of medicine requires the quantification of effect modification by different risk groups, rather than effects averaged over multiple risk groups. Further, the required subgroup analyses necessitate a sample size that would make most randomized trials prohibitively expensive. Therefore, we should seriously consider making a start at identifying and measuring all possible effect modifiers in observational research, or risk never making any meaningful progress in the personalization of medicine. The obvious limitation in our “call to observational clinical research” lies

in the effectiveness of the proposed approach to really deal with all relevant confounding by indication and provide us with valid estimates of the effects of treatments.

To validly estimate the effects of treatment, we need to establish conditional exchangeability.^{17–19} We will arguably never be entirely sure whether we achieved this goal or not.^{17–19} However, this is in its essence not different from a randomized trial, in which we assume all prognostic factors to be equally distributed by chance, but get no guarantee.²⁴ This is the reason we always include a “Table 1” with all important prognostic indicators. We know that, by chance alone, the prognosis can be different between the treated and untreated, and there will not be exchangeability. We can only be sure of the absence of confounding by coincidence over an infinite number of trials, which again is a theoretical impossibility.²⁴ There are two main considerations in achieving conditional exchangeability in observational research. First, have we identified and accurately measured all relevant confounders? Second, have we modeled them correctly?

The first consideration we have briefly hinted at already. In many fields of epidemiology, confounding can never be expected to be solved completely, due to the presence of unmeasured or unmeasurable confounders.^{25,26} However, in clinical epidemiology, confounding is caused by a physician’s decision to treat a patient. Therefore, it is theoretically impossible to have truly unknown and unknowable confounders. The physician has to know; otherwise the variable can never influence the decision. Unmeasured genetic variants, for example, could be effect modifiers and interesting targets for further research, but are not a concern as possible confounders, because they cannot influence the physician’s decision. By the same token, measurement error cannot result in residual confounding. As long as we assure that we correct with the same measurement error that the physician used in his or her decision, we are still correcting completely for the confounding that the physician introduced. If the physician knows, it should be possible, with the help of that physician, to record the confounding variable with the same accuracy the physician used.

The second consideration has to do with the modeled form of the relation among the confounder, the treatment, and the outcome.^{27,28} We have not discussed this issue yet and have assumed statistical correction to be perfect throughout. This is a simplification of reality, but not one that could invalidate our conclusion that more observational research is needed. To be able to more accurately correct for a confounder, more data are needed. This is actually the main reason we consider observational research preferable for the quantification of effect modification: in observational research, it is easier (ie, cheaper) to gather more data.

In conclusion, the personalization of medical decision-making could greatly be enhanced by using more observational clinical research. However, great care should always be taken with respect to confounding by indication, and conclusions from observational research should never be accepted lightly. Appropriate appreciation of the evidence provided by different observational studies might require much more active discussion in the medical literature, instead of the current, often limited interaction between authors and readers.

Author contributions

RAM and JGvdB were responsible for conceptualization, drafting, writing, editing, revising, and intellectual contributions. MSA and JGM were responsible for editing, revising, and intellectual contributions. All authors contributed toward data analysis, drafting and revising the paper and agree to be accountable for all aspects of the work.

Disclosure

Related authors: RAM and JGM are married to each other. The other authors report no other conflicts of interest in this work.

References

- Sibbald B, Roland M. Understanding controlled trials. Why are randomised controlled trials important? *BMJ*. 1998;316(7126):201.
- Rubin DB. Bayesian inference for causal effects: the role of randomisation. *Ann Stat*. 1978;6(1):34–58.
- Bosco JL, Silliman RA, Thwin SS, et al. A most stubborn bias: no adjustment method fully resolves confounding by indication in observational studies. *J Clin Epidemiol*. 2010;63(1):64–74.
- Freemantle N, Marston L, Walters K, Wood J, Reynolds MR, Petersen I. Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. *BMJ*. 2013;347:f6409.
- Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992;268(17):2420–2425.
- Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71–72.
- Rothwell PM. Can overall results of clinical trials be applied to all patients? *Lancet*. 1995;345(8965):1616–1619.
- Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet*. 2005;365(9454):176–186.
- Rothwell PM, Mehta Z, Howard SC, Gutnikov SA, Warlow CP. Treating individuals 3: from subgroups to individuals: general principles and the example of carotid endarterectomy. *Lancet*. 2005;365(9455):256–265.
- Rothwell PM. External validity of randomised controlled trials: “to whom do the results of this trial apply?” *Lancet*. 2005;365(9453):82–93.
- Kent DM, Nelson J, Dahabreh IJ, Rothwell PM, Altman DG, Hayward RA. Risk and treatment effect heterogeneity: re-analysis of individual participant data from 32 large clinical trials. *Int J Epidemiol*. 2016;45(6):2075–2088.
- Signorello LB, McLaughlin JK, Lipworth L, Friis S, Sørensen HT, Blot WJ. Confounding by indication in epidemiologic studies of commonly used analgesics. *Am J Ther*. 2002;9(3):199–205.
- Middelburg RA, van de Watering LM, van der Bom JG. Blood transfusions: good or bad? Confounding by indication, an underestimated problem in clinical transfusion research. *Transfusion*. 2010;50(6):1181–1183.
- Strom BL, Miettinen OS, Melmon KL. Postmarketing studies of drug efficacy: when must they be randomized? *Clin Pharmacol Ther*. 1983;34(1):1–6.
- Strom BL. Methodologic challenges to studying patient safety and comparative effectiveness. *Med Care*. 2007;45(10 Suppl 2):S13–S15.
- Kyriacou DN, Lewis RJ. Confounding by indication in clinical research. *JAMA*. 2016;316(17):1818–1819.
- Greenland S, Robins JM. Identifiability, exchangeability, and epidemiological confounding. *Int J Epidemiol*. 1986;15(3):413–419.
- Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu Rev Public Health*. 2000;21:121121–145145.
- Hernán MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health*. 2004;58(4):265–271.
- Vandenbroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology: the need for a pluralistic approach. *Int J Epidemiol*. 2016;45(6):1776–1786.
- Vincent JL. We should abandon randomized controlled trials in the intensive care unit. *Crit Care Med*. 2010;38(10 Suppl):S534–S538.
- Concato J. Study design and “evidence” in patient-oriented research. *Am J Respir Crit Care Med*. 2013;187(11):1167–1172.
- Concato J. When to randomize, or “Evidence-based medicine needs Medicine-based evidence.” *Pharmacoepidemiol Drug Saf*. 2012;21(Suppl 2):6–12.
- Rothman KJ. Epidemiologic methods in clinical trials. *Cancer*. 1977;39(4 Suppl):1771–1775.
- Uddin MJ, Groenwold RH, Ali MS, et al. Methods to control for unmeasured confounding in pharmacoepidemiology: an overview. *Int J Clin Pharm*. 2016;38(3):714–723.
- Vandenbroucke JP. When are observational studies as credible as randomised trials? *Lancet*. 2004;363(9422):1728–1731.
- Brenner H, Blettner M. Controlling for continuous confounders in epidemiologic research. *Epidemiology*. 1997;8(4):429–434.
- Becher H. The concept of residual confounding in regression models and some applications. *Stat Med*. 1992;11(13):1747–1758.

Clinical Epidemiology

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventive initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification,

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>

Dovepress

systematic reviews, risk and safety of medical interventions, epidemiology and biostatistical methods, and evaluation of guidelines, translational medicine, health policies and economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.