Treatment-Line Versus Patient-Level Matching: A Case Study in Oncology

Xavier Pouwels, MSc, Bram Ramaekers, PhD; Manuela Joore, PhD, Maastricht University Medical Centre+, Maastricht, The Netherlands

There is an alternative way of performing matching when using data collected from daily clinical practice. Treatment-line matching influences model inputs and results, which ultimately may affect reimbursement decisions.

Observational data, confounding by indication, and matching methods

Observational data are increasingly used to inform economic evaluations informing healthcare decision making. However, the biggest threat when using observational data to compare 2 (or more) treatments is the lack of randomization (ie, the comparison is subject to selection bias due to confounding by indication). This means that patients in the treatment groups may have different baseline characteristics that may be related to treatment assignment and the outcome of interest. In those cases, naïvely comparing treatment groups will most likely result in a biased estimate of the treatment's cost-effectiveness. Statistical methods, including regression-based adjustments, matching, and instrumental variables methods, have been developed to address this issue.[1]

Matching methods (eg, propensity score matching and genetic matching) are the main subject of the current article. These methods aim at increasing the similarity in observed baseline characteristics between patients in the intervention and comparator groups. When using propensity score matching, the probability of being assigned to treatment is estimated per patient based on observed baseline characteristics (potential confounders). This probability is then used to match comparator patients with the most similar baseline characteristics to patients in the intervention group.[2] Genetic matching is a search algorithm that automatically maximizes the similarity in prespecified baseline characteristics between the intervention and comparator groups.[3]

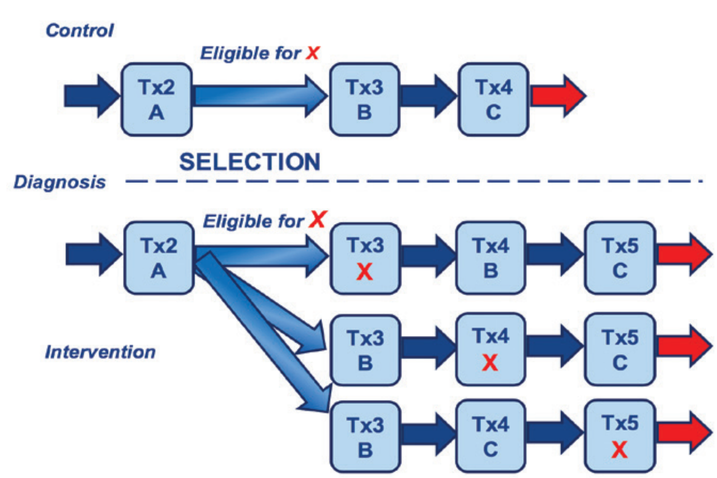
Why treatment-line matching?

In multiple disease areas such as oncology, rheumatoid arthritis, and cardiovascular disorders, patients typically receive multiple treatment lines, which may create 2 issues. First, comparator patients are usually included in the comparator group at the moment they become eligible for the intervention. This creates an imbalance between the comparator and the intervention group, if (a proportion of) the patients in the intervention group received the intervention later in the treatment pathway than at the moment they became eligible for it. In this situation, patients in the intervention group may be more heavily pretreated than patients included in the comparator group. This imbalance in pretreatment may consequently influence the cost-effectiveness of the intervention versus the comparator. This is illustrated in Figure 1.

Second, the performance of matching methods is influenced by the overlap in baseline characteristics between patients

Figure 1. Imbalance in pretreatment between control and intervention groups.

This illustrates the imbalance in pretreatment between patients who are identified based on the eligibility criteria for the intervention (ie, control), and patients who may receive the intervention later in the treatment pathway (ie, intervention).



Tx indicates treatment.

Figure 2. Distribution of baseline age based on the genetic matched populations.

The illustration represents the distribution of baseline age in the intervention group (y-axis) versus the usual care group (x-axis). Ideally, all dots should lie on the line, which would indicate a perfect overlap of baseline characteristics between the groups.

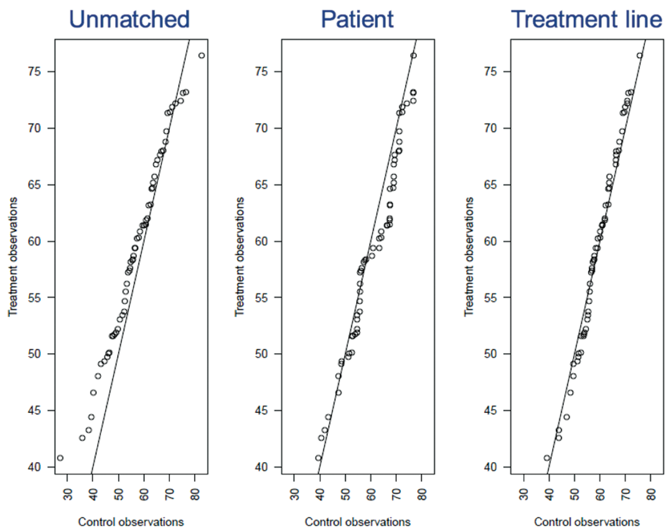
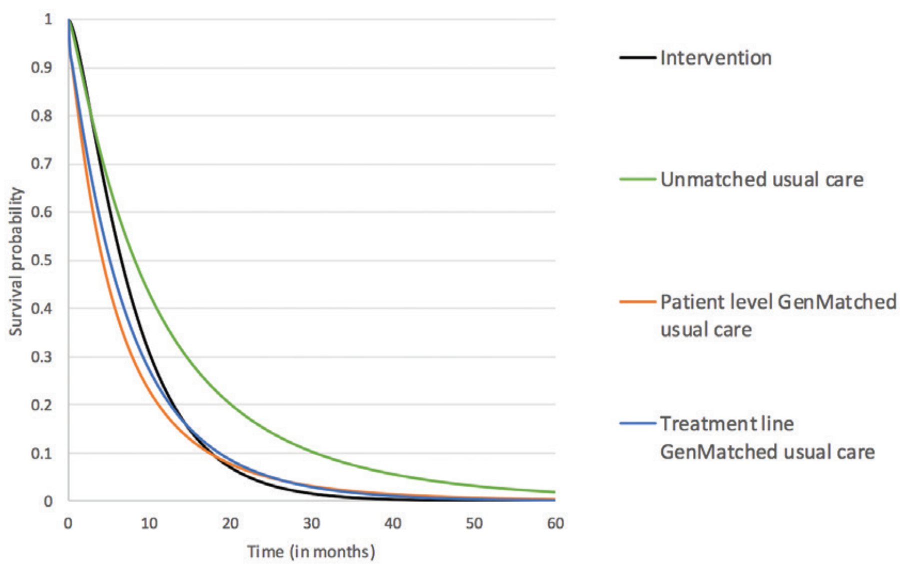


Figure 3: Overall survival curves used as model inputs for each group (based on a generalized gamma distribution).



GenMatched indicates matched using genetic matching.

in the intervention and comparator groups and the size of the comparator group. In case of poor overlap in baseline characteristics and small number of comparator patients, matching methods may not be able to increase the similarity in baseline characteristics adequately between the intervention and comparator groups. Additionally, the variance surrounding baseline characteristics in the comparator group may be underestimated.

By considering all treatment lines administered to comparator patients as an individual comparator, the number of potential comparators is increased and the fact that patients do not receive the intervention when they become eligible for it is reflected in the comparator group. In other words, including treatment lines in the pool of potential comparators results in including different "versions" of the comparator patients in the pool of comparator. This process may be related to matching with replacement, where comparator patients may be included multiple times in the comparator group.

An illustration in oncology

The current case concerns an economic evaluation of an oncology treatment (the intervention) versus a comparator (usual care). This analysis was based on data collected in daily clinical practice, and the comparison is consequently subject to confounding by indication. Hence, we decided to apply genetic matching to obtain a usual care group that was similar to the intervention group. Since patients in the intervention group had often received the intervention later in the treatment pathway than when they became eligible for it, treatment-line matching might be indicated. We decided to apply both patient-level matching and treatment-line matching to investigate whether treatment-line matching would indeed increase the similarity in baseline characteristics and what the influence of treatment-line matching would be on the results. For completeness, a comparison with the unmatched usual care group was performed. This analysis therefore contains 3 comparisons: (1) intervention versus unmatched usual care, (2) intervention versus patient-level–matched usual care, and (3) intervention versus treatment-line-matched usual care.

The cost-effectiveness model was a 3 health states (progression-free, progressed disease, and death), partitioned survival model. Patients entered the model in the progression-free health state and could either progress or die. Patients in the progressed-disease health state could not transition to the progression-free health state. Effectiveness and resource use estimates were obtained from the database, while utilities and prices were obtained from the literature. In the cost-effectiveness model, progression-free survival and overall survival were estimated through parametric time-to-event models.

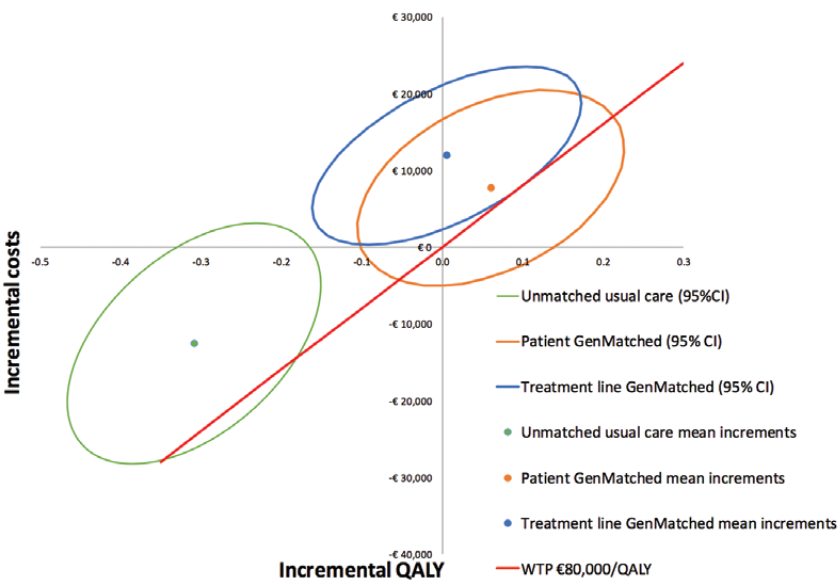
In total, there were 90 patients who received the intervention and 321 patients who composed the unmatched usual care group. The 2 matched usual care groups were composed of 90 patients (or treatment lines) each. When analyzing the similarity in baseline characteristics, based on visual inspection of eQQ plots (Figure 2) and a statistical criterion (the bootstrapped Kolmogorev-Smirnov test), we observed that the treatment-line-matched groups were, in general, more similar to the intervention group. >

Treatment-line matching influenced the effectiveness of usual care. The treatment-line usual care group had the longest overall survival estimates than the patient-level–matched usual care group. The unmatched usual care group had the longest survival compared to the matched usual care groups (Figure 3). The unmatched usual care group also had the highest costs associated with the progression-free and progressed-disease health states. The treatment-line-matched usual care group had the longest progressed-disease health state costs than the patient-level–matched usual care group.

These differences in survival and health state costs resulted in differences in total quality-adjusted life years (QALY) gain and total costs obtained by the different usual care groups. The cost-effectiveness of the intervention versus the usual groups was thus affected by whether matching was performed on patients or treatment lines. The intervention was dominated by usual care when compared to the unmatched usual care group, but was more effective and more costly than the treatment-line-matched and patient-level–matched usual care groups (Figure 4).

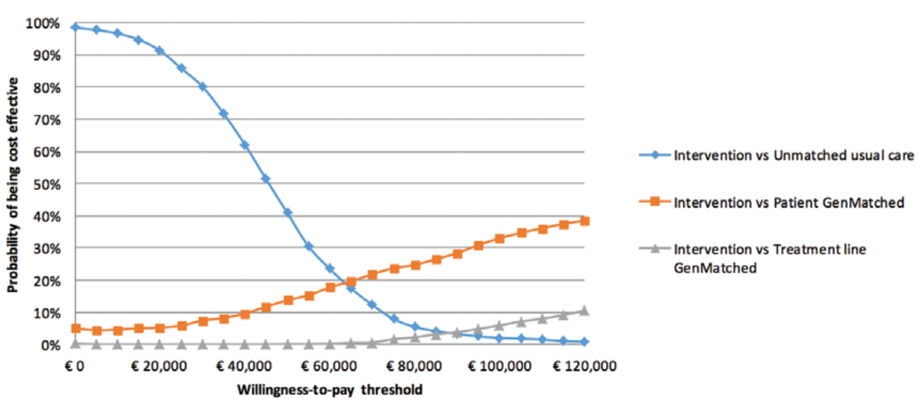
The uncertainty surrounding the results of the comparison of the intervention versus the treatment-line-matched usual care group was lower than the uncertainty surrounding the results of the comparison of the intervention versus the patient-level-matched usual care group. Finally, this resulted in different probabilities of the intervention being cost-effective in each comparison with the usual care groups (Figure 5).

Figure 4: Incremental costs and QALY for each comparison.



CI indicates confidence interval; GenMatched, matched using genetic matching; QALY, quality-adjusted life year; WTP, willingness to pay.

Figure 5: Cost-effectiveness acceptability curve of each comparison.



CI indicates confidence interval; GenMatched, matched using genetic matching.

Conclusions

Through this short article, we hope to raise the awareness concerning the possibility of using matching methods on treatment-lines. This case study demonstrates that treatment-line matching improved the similarity in baseline characteristics between the intervention and usual care groups compared to patient-level matching. Treatment-line matching also influenced the model inputs, results, and the uncertainty surrounding the results, which may affect reimbursement decisions.

NOTE: The empirical data in this article has been systematically modified.

References

1. Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci.* 2010;25(1):1-21. doi:10.1214/09-sts313.

2. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-55. doi:10.1093/biomet/70.1.41.

3. Sekhon JS, Grieve RD. A matching method for improving covariate balance in cost-effectiveness analyses. *Health Econ*. 2012;21(6):695-714. doi:10.1002/hec.1748.