Appendix

Average treatment effect on treated

Following Deheja and Wahba (1), we denoted the outcome (in this case, whether an individual will quit their job) of individual \( i \) as \( Y_{i1} \) if they were diagnosed with cancer, and \( Y_{i0} \) if they were not diagnosed with cancer. Hereafter, we refer to those who were diagnosed with cancer as the ‘treatment’ group. The treatment effect on a single entity, \( \tau_i \) is defined as \( \tau_i \equiv Y_{i1} - Y_{i0} \). We aimed to estimate the average treatment effect on the treatment group (ATT); that is,

\[
(S1.1) \quad ATT = E(Y_{i1}|T = 1) - E(Y_{i0}|T = 1)
\]

\[
= [E(Y_{i1}|T = 1) - E(Y_{i0}|T = 0)] + [E(Y_{i0}|T = 0) - E(Y_{i0}|T = 1)],
\]

where \( T = 1, \) or \( 0 \) indicates the \( i \)th individual was assigned to the treated (control) group. If and only if the second term was zero, we could obtain a consistent and unbiased estimate of true ATT by simply comparing the means of the two groups. However, in our case, this assumption was unrealistic because cancer patients can potentially engage in unhealthy behaviours before diagnosis and, therefore, their counter-factual outcomes can differ from those of non-cancer patients.

To formalise the PSM estimator, we assumed several assumptions (observable covariates for person \( i \) are denoted as \( X_i \)).

[Assumption 1] (Unconfoundedness for controls)

\[
(S1.2) \quad Y_{i0} \perp T_i \mid X_i \Rightarrow E(Y_{i0}|X_i, T_i) = E(Y_{i0}|X_i)
\]

[Assumption 2] (Overlap)

\[
(S1.3) \quad \Pr(T = 1 \mid X_i) < 1
\]

These two assumptions are a weaker version of the ‘strong ignorability’ defined by Rosenbaum and Rubin (2). Under these assumptions, ATT conditional on covariates, \( ATT(X) \), was

\[
(S1.4) \quad ATT = E(Y_{i1}|X_i, T = 1) - E(Y_{i0}|X_i, T = 1)
\]

\[
= E(Y_{i1}|X_i, T = 1) - E(Y_{i0}|X_i, T = 0).
\]
In the above equation, we could estimate ATT(X) by data. However, as highlighted in Caliendo and Kopeinig (3), ‘conditioning on all relevant covariates is limited in the case of a high dimensional vector X.’ To address this problem, we utilised the procedure in Rosenbaum and Rubin (2) and conditioned the probability of treatment as a function of covariates X, which is called ‘propensity score’. If the abovementioned assumptions were satisfied, $Y_{i0}$ and $T_i$ were independent conditional on the propensity score; that is, $Y_{i0} \perp T_i \mid p(X_i)$ (Lemma 2.14 in reference (4)). Then, we constructed the ATT conditional on the propensity score, PSM, as follows:

$$\tau_{PSM} = E_{p(X)} \mid T = 1 \left[ E(Y_{1i} \mid T = 1, p(X)) - E(Y_{0i} \mid T = 0, p(X)) \right],$$

where $E_{p(X)} \mid T = 1$ denotes the expectation of the treated population with regard to the propensity score, which implicitly claims the common support restriction.

**Nearest Neighbour Matching**

We employed nearest neighbour matching, in which a cancer patient was matched with a weighted average of the first 15 closest individuals in the control group (NN15). In this case, ATT could be obtained using the following formula (5):

$$\overline{\tau}_{PSM} = \frac{1}{|T|} \sum_{i \in T} \left[ Y_i^T - \sum_{j \in C(i)} w_{ij} Y_j^C \right],$$

where $T$ denotes treatment group, $|T|$ is the number of entities in the treatment group, $C(i)$ indicates the subset of a control group matched with the treated $i$, and $w_{ij}$ is weight on the $j$th individual. For detailed discussion regarding weighting strategies, see Heckman, Ichimura, and Todd (6).

In order to avoid bad matching, we set the following calliper restriction:

$$j \text{ is included in } C(i) \text{ if } p_j \text{ is one of the 15 closest scores to } p_i \text{ and } ||p_i - p_j|| < c \text{ excluded otherwise}$$
As we note in the paper, we set $c = 0.01$. We computed the standard error for $\hat{\tau}_{PSM}$ using bootstrapping with 200 replications.

**Covariate Balancing Score**

The covariate balancing scores before and after PSM in S1 Table.
### S1 Table. Covariate balancing scores.

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>Lagged</td>
<td>Simultaneous</td>
</tr>
<tr>
<td></td>
<td>U</td>
<td>M</td>
</tr>
<tr>
<td>LR $\chi^2$</td>
<td>108.69</td>
<td>1.20</td>
</tr>
<tr>
<td>p-value</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Mean% Bias</td>
<td>7.5</td>
<td>1.1</td>
</tr>
<tr>
<td>Med% Bias</td>
<td>5.5</td>
<td>0.8</td>
</tr>
<tr>
<td>Rubin’s B</td>
<td>58.0*</td>
<td>8.4</td>
</tr>
<tr>
<td>Rubin’s R</td>
<td>0.84</td>
<td>0.99</td>
</tr>
</tbody>
</table>

**Note:** ‘LR $\chi^2$’ indicates likelihood-ratio statistics of the joint insignificance of all covariates. ‘P-value’ indicates the corresponding p-values of the tests. Mean (Median)% bias is a summary indicator of the distribution of the standardised differences of the covariates after matching; generally, maximum bias at 10% is considered to signify negligible difference [11]. ‘Rubin’s B’ is the absolute standardised difference of the means of the linear index of the propensity score in the treatment group and matched controls. ‘Rubin’s R’ is the ratio of the treatment group to matched non-treated group in terms of the variances of the propensity score index. For sufficient balances, Rubin recommends that B be less than 25% and that R be between 0.5 and 2%. * if B > 25, R < [0.5; 2] [12,13].
Reference


