Marginal Structural Models as a Tool for Standardization

Tosiya Sato and Yutaka Matsuyama

Abstract: In this article, we show the general relation between standardization methods and marginal structural models. Standardization has been recognized as a method to control confounding and to estimate causal parameters of interest. Because standardization requires stratification by confounders, the sparse-data problem will occur when stratified by many confounders and one then might have an unstable estimator. A new class of causal models called marginal structural models has recently been proposed. In marginal structural models, the parameters are consistently estimated by the inverse-probability-of-treatment weighting method. Marginal structural models give a nonparametric standardization using the total group (exposed and unexposed) as the standard. In epidemiologic analysis, it is also important to know the change in the average risk of the exposed (or the unexposed) subgroup produced by exposure, which corresponds to the exposed (or the unexposed) group as the standard. We propose modifications of the weights in the marginal structural models, which give the nonparametric estimation of standardized parameters. With the proposed weights, we can use the marginal structural models as a useful tool for the nonparametric multivariate standardization.

Key Words: causal models, confounding, epidemiologic methods, standardization, weighted analysis

(Epidemiology 2003;14: 680–686)
negative: LYM = 0). No effect of tamoxifen was observed in the crude data (the crude risk ratio is 1.01). However, possible protective effects and effect-measure modification were observed in the stratified data (the stratum-specific risk ratios are 0.91 and 0.67, respectively).

When the target population of the study is the exposed group and one needs to summarize the overall effect of exposure in the ratio scale, the standardized mortality/morbidity ratio (SMR) is such a measure, which is given by:

$$\text{SMR} = \frac{\sum_k x_k}{\sum_k n_k} \times \frac{\sum_k y_k}{\sum_k m_k}$$

The numerator of SMR can be interpreted as the number of deaths (or events) in the exposed group when the exposed group was actually exposed, which is identical to the observed number of deaths. In contrast, the denominator can be interpreted as the expected number of deaths in the exposed group if the exposed group had not been exposed.\(^4,6\) Note that this expectation is not the null (no effect) expectation, which is given by \(\sum_k n_k t_k / N_k\), where \(t_k = x_k + y_k\) and \(N_k = n_k + m_k\). It is instead the counterfactual number of deaths as if the exposed had not been exposed, because in the \(k\)th stratum we could substitute \(y_k/m_k\) for the counterfactual disease frequency in the exposed when confounding is removed by stratification.\(^6\)

In the example in Table 2, SMR is calculated by:

$$\text{SMR} = \frac{368 + 96}{1215 \times 760} \times \frac{1334 \times 171}{1592} = \frac{368 + 96}{404.5 + 143.3} = \frac{464}{547.8} = 0.85.$$  

The SMR is interpreted as the proportionate change in risk (or rate) of the exposed group produced by exposure. In the calculation of SMR, the exposed group is used as the standard population. This is also known as the indirect standardization. Other choices of target population that are commonly used in epidemiologic studies are the unexposed group and the total group (combining the exposed and the unexposed groups).\(^13\)

When the unexposed group is the target population and is chosen as the standard population, the standardized risk (or rate) ratio in the unexposed is given by:

$$\text{SRR}_U = \frac{\sum_k x_k}{\sum_k n_k} \times \frac{\sum_k y_k}{\sum_k m_k}$$

This is known as the direct standardization. In our example, it is calculated as:

$$\text{SRR}_U = \frac{760 \times 368}{1215 + 1592 \times 96} = \frac{230.2 + 114.6}{253 + 171} = \frac{344.8}{424} = 0.81.$$  

### Table 1. Notation for the \(k\)th Stratum

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Exposed</th>
<th>Unexposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events</td>
<td>(x_k)</td>
<td>(y_k)</td>
</tr>
<tr>
<td>Denominator</td>
<td>(n_k)</td>
<td>(m_k)</td>
</tr>
</tbody>
</table>

### Table 2. Crude and Stratified Count Data of the Tamoxifen Use and the Recurrence of Breast Cancer

<table>
<thead>
<tr>
<th></th>
<th>TAM = 1</th>
<th>TAM = 0</th>
<th>TAM = 1</th>
<th>TAM = 0</th>
<th>TAM = 1</th>
<th>TAM = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>RECUR = 1</td>
<td>464</td>
<td>424</td>
<td>368</td>
<td>253</td>
<td>96</td>
<td>171</td>
</tr>
<tr>
<td>Total no.</td>
<td>2549</td>
<td>2352</td>
<td>1215</td>
<td>760</td>
<td>1334</td>
<td>1592</td>
</tr>
<tr>
<td>Risk ratio</td>
<td>1.01</td>
<td>0.91</td>
<td>0.67</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

LYM = 1 (positive for lymph node metastasis at surgery), 0 (negative); TAM = 1 (exposed to tamoxifen), 0 (unexposed); RECUR = 1 (recurrence of breast cancer).

© 2003 Lippincott Williams & Wilkins

---

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
SRR is interpreted as the proportionate change in the risk (or rate) that would have occurred in the unexposed group had they been exposed.

When the total group (combining the exposed and the unexposed) is the target population, the standardized risk (or rate) ratio in the total group,

\[ SRR_T = \frac{\sum x_k / n_k}{\sum y_k / m_k}, \]

is calculated as:

\[ SRR_T = \frac{1975 \times 368 + 2926 \times 96}{1975 \times 253 + 2926 \times 171} = \frac{598.2 + 210.6}{657.5 + 314.3} = \frac{808.8}{971.8} = 0.83 \]

SRR is interpreted as the proportionate change in risk (or rate) in the total group under complete exposure and complete nonexposure.

The interpretations of these standardized parameters are still valid even when there is effect–measure modification, in other words, stratum specific ratios are heterogeneous.

### Marginal Structural Models and Standardization

We explain the marginal structural models in the context of causal risk ratio estimation. Application to the estimation of other effect measures is straightforward. In the estimation of the effect of a dichotomous exposure E (1: exposed, 0: unexposed) on a dichotomous outcome D (1: event occurred, 0: no event), we consider the contrast of the following potentially counterfactual probabilities:

\[ P(D = 1|\text{set } E = 1) \text{ and } P(D = 1|\text{set } E = 0) \]

for the probability of D = 1 if everyone in the target population had been exposed ("set E = 1") and if everyone in the target population had not been exposed ("set E = 0").

Note that the counterfactual probability, \( P(D = 1|\text{set } E = e) \), can be different from the observed probability, \( P(D = 1|E = e) \), because the latter refers to a subset of population members with \( E = e \), whereas the former refers to all population members.

In the marginal structural models, the marginal distribution of the counterfactual probabilities is modeled as:

\[ \log P(D = 1|\text{set } E = e) = \alpha_0 + \alpha_1 e, \]  
\[ \log P(D = 1|E = e) = \theta_0 + \theta_1 e, \]

where \( \exp(\alpha_1) \) is the causal risk ratio. The corresponding log-linear model for the observed probability can be written as:

\[ \log P(D = 1|E = e) = \theta_0 + \theta_1 e, \]

The interpretations of these standardized parameters are still valid even when there is effect–measure modification, in other words, stratum specific ratios are heterogeneous.
Table 3: Inverse Probability of Treatment Weights $w_T$, SMR Weights $w_E$, and Composition of Pseudopopulation

| LYM | TAM | RECUR | Observed No. | P(E|Z) | $w_T$ | Pseudo No. | Total | $w_E$ | Pseudo No. Exposed |
|-----|-----|-------|--------------|-------|-------|------------|-------|-------|-------------------|
| 1   | 1   | 1     | 368          | 0.615 | 1.63  | 598.2      | 1     | 368   |
| 1   | 1   | 0     | 847          | 0.615 | 1.63  | 1376.8     | 1     | 847   |
| 1   | 0   | 1     | 253          | 0.385 | 2.60  | 657.5      | 1.60  | 404.5 |
| 1   | 0   | 0     | 507          | 0.385 | 2.60  | 1317.5     | 1.60  | 810.5 |
| 0   | 1   | 1     | 96           | 0.456 | 2.19  | 210.6      | 1     | 96    |
| 0   | 1   | 0     | 1238         | 0.456 | 2.19  | 2715.4     | 1     | 1238  |
| 0   | 0   | 1     | 171          | 0.544 | 1.84  | 314.3      | 0.84  | 143.3 |
| 0   | 0   | 0     | 1421         | 0.544 | 1.84  | 2611.7     | 0.84  | 1190.7|

LYM = 1 (positive for lymph node metastasis at surgery), 0 (negative); TAM = 1 (exposed to tamoxifen), 0 (unexposed); RECUR = 1 (recurrence of breast cancer), 0 (no recurrence).

Table 4: Crude Data From the Pseudopopulation Created by the Inverse Probability of Treatment Weights

<table>
<thead>
<tr>
<th></th>
<th>Exposed</th>
<th>Unexposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recurrence</td>
<td>808.8</td>
<td>971.8</td>
</tr>
<tr>
<td>No. of women</td>
<td>4901</td>
<td>4901</td>
</tr>
<tr>
<td>Crude risk ratio = 0.83</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

extensions of SMR or $SRR_U$ in the marginal structural models. Similar to $SRR_T$, SMR can be rewritten as:

$$SMR = \frac{\sum x_k}{\sum y_k} = \frac{\sum x_k \left( \frac{n_k}{N_k} \right)^{-1} n_k}{\sum y_k \left( \frac{m_i}{N_k} \right)^{-1} n_k}$$

The weights implicitly used in $SMR$ can be interpreted as the inverse of the conditional probability of receiving the subject’s own exposure multiplied by the conditional probability of receiving exposure regardless of the subject’s actual exposure status (see Appendix 1). This leads to a new weight for the $i$th subject:

$$w_{Ei} = \frac{P(E = 1|Z = z_i)}{P(E = e|Z = z_i)}$$

(3)

We call this the SMR weight. In the SMR weight, the denominator works to control confounding in the same way as in the inverse probability of treatment weights; the numerator reweights the pseudo-population to give it the distribution of covariates in the target population (here, the exposed).

The weighted analysis of the association model (2) with new weights (3) gives unbiased estimates of the parameters in the following marginal structural model:

$$\log P(D = 1|\text{set } E = e, E = 1) = \beta_0 + \beta_1 e,$$

where “set $E = e, E = 1$” means what would have occurred in the exposed group had the exposed group taken the exposure $e$. The resultant estimator for $\exp(\theta_1)$, which is
TABLE 5. Crude Data from the Pseudopopulation Created by the SMR Weights

<table>
<thead>
<tr>
<th></th>
<th>Exposed</th>
<th>Unexposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recurrence</td>
<td>464</td>
<td>547.8</td>
</tr>
<tr>
<td>No. of women</td>
<td>2549</td>
<td>2549</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crude risk ratio = 0.85</td>
</tr>
</tbody>
</table>

TABLE 7. Adjuvant Tamoxifen Use and Recurrence of Breast Cancer

<table>
<thead>
<tr>
<th></th>
<th>Exposed</th>
<th>Unexposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recurrence</td>
<td>464</td>
<td>424</td>
</tr>
<tr>
<td>Women-years</td>
<td>17228</td>
<td>17461</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Estimated rate ratio = 1.11 (95% confidence interval = 0.97-1.27)</td>
</tr>
</tbody>
</table>

consistent with causal parameter \( \exp(\beta_i) \), has the interpretation as a nonparametric multivariate extension of the standardized mortality ratio. For the exposed subjects the weight \( w_{EI} \) is always 1 and for the unexposed subjects it is the conditional exposure odds. Table 3 also displays the SMR weighting scheme. The SMR weight \( w_E \) in the LYM = 1 stratum is calculated as the total number of exposed women (1215) divided by the total number of unexposed women (760), which yields 1.60. The last column, “Pseudo No. Exposed,” represents the number of women in the SMR-weighted pseudo-population. Table 5 displays the crude data from the pseudo-population created by the SMR weights. The crude risk ratio is 0.85, which is identical to SMR.

Parallel arguments give the extension of the standardization method with the unexposed group as the standard. The marginal structural model can be written as:

\[
\log P(D = 1|E = e, E = 0) = \gamma_0 + \gamma_1 e,
\]

and we have unbiased estimates of \( (\gamma_0, \gamma_1) \) by the weighted analysis of the association model (2) with weights

\[
w_{ui} = \frac{P(E = 0|Z = z_i)}{P(E = e|Z = z_i)}. \tag{4}
\]

In Table 6, a summary of different sets of weights is given when we use the logistic model

\[
\logit P(E = 1|Z = z_i) = \lambda' z_i
\]

in the estimation of propensity scores. Here we include a constant 1 in \( z_i \) so that the logistic model can contain the intercept.

**Tamoxifen Use and Recurrence of Breast Cancer**

Table 7 shows crude person-time data from an observational study of the effect of postoperative tamoxifen on the recurrence of breast cancer.\(^1\)\(^4\) The study subjects were 6148 women who had been diagnosed with unilateral primary breast cancer and who had received surgical treatment between 1982 and 1990 at 9 institutions in Japan. The study was initiated in 1995 and the end of follow up was March 1996. Information on each patient was retrospectively obtained from medical records or a prospectively compiled computer database at each institution. Among 6148 women, information on recurrence was obtained for 4901 women shown in Table 7 as well as Table 2. The crude rate ratio was 1.11 with a 95% confidence interval of 0.97-1.27. There was no protective effect of adjuvant tamoxifen use on the recurrence of breast cancer. However, because the exposure to adjuvant tamoxifen use was not randomized, the crude results might have been confounded as in the previous risk-ratio analyses.

We conducted weighted Poisson regression analysis\(^1\) in which the recurrence rate \( r \) was modeled as:

\[
\log r = \omega_0 + \omega_1 e \tag{5}
\]

with the estimated SMR weights \( \hat{w}_{EI} \) and where \( e \) is the exposure indicator (1, exposed to tamoxifen; 0, unexposed). The propensity score for each subject \( i \) was estimated by the

\[
\logit P(E = 1|Z = z_i) = \lambda' z_i
\]

\( \hat{\lambda} \) are corresponding maximum likelihood estimates to \( \lambda \).

TABLE 6. Weights Used in Marginal Structural Models With Different Standard Populations When Propensity Scores are Estimated by the Logistic Model: \( \logit P(E = 1|Z = z_i) = \lambda' z_i \)

<table>
<thead>
<tr>
<th>Standard Population</th>
<th>Exposed</th>
<th>Unexposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>( 1 + \exp(-\hat{\lambda} z_i) )</td>
<td>( 1 + \exp(\hat{\lambda} z_i) )</td>
</tr>
<tr>
<td>Exposed</td>
<td>( 1 )</td>
<td>( \exp(\hat{\lambda} z_i) )</td>
</tr>
<tr>
<td>Unexposed</td>
<td>( \exp(-\hat{\lambda} z_i) )</td>
<td>( 1 )</td>
</tr>
</tbody>
</table>

\( \hat{\lambda} \) are corresponding maximum likelihood estimates to \( \lambda \).

TABLE 8. Crude Data From the Pseudopopulation Created by the SMR Weighting

<table>
<thead>
<tr>
<th></th>
<th>Exposed</th>
<th>Unexposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recurrence</td>
<td>464</td>
<td>554.0</td>
</tr>
<tr>
<td>Women-years</td>
<td>17228</td>
<td>18303.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Estimated rate ratio = 0.89</td>
</tr>
</tbody>
</table>

(Robust 95% confidence interval = 0.78-1.02)
The marginal structural models with the inverse probability of treatment weights give the nonparametric multivariate extension of the standardization method with the total group as the standard population. We have shown that other choices of weights give the extension of the standardization method with other groups as the standard. All the weights can be calculated using the estimation of the propensity scores (Table 6). We primarily considered the estimation of ratio measures in the previous sections. It is straightforward to apply the proposed weights to the estimation of difference measures. Marginal structural models with various sets of weights can be used as a useful tool for the standardization methods, depending on the appropriate choice of target population.

Appendix 1

Under the assumption of no unmeasured confounder, the counterfactual probability of any target population, \(P(D = 1|\text{set } E = e, \text{target population})\), is a weighted average of the stratum-specific risks with weights proportional to the distributions of \(Z\) in the target population;

\[
P(D = 1|\text{set } E = e, \text{target population}) = \sum_z P(D = 1|E = e, Z = z) P(Z = z|\text{target population}),
\]

where the sum is over the possible values of \(Z\). Simple algebra gives

\[
\sum_z P(D = 1|E = e, Z = z) P(Z = z|\text{target population}) = \sum_z P(D = 1|E = e, Z = z)
\]

\[
\times P(Z = z|\text{target population})
\]

\[
= \sum_z \frac{P(D = 1, E = e, Z = z)}{P(E = e|Z = z)} \times \frac{P(Z = z|\text{target population})}{P(Z = z)}.
\]
where \( P(Z = z) \) is the distributions of \( Z \) in the entire study (total) population. The target population might sometimes be completely external.

When the total population is the target population, \( P(Z = z | \text{target population}) = P(Z = z) \) and we have

\[
P(D = 1 | \text{set } E = e) = \sum_z \frac{P(D = 1, E = e, Z = z)}{P(E = e | Z = z)}.
\]

This leads to the inverse probability of treatment weighting. Similarly, when the exposed group is the target population, \( P(Z = z | \text{target population}) = P(Z = z | E = 1) \) and

\[
P(D = 1 | \text{set } E = e, E = 1) = \frac{1}{P(E = 1)} \sum_z P(D = 1, E = e, Z = z) \frac{P(E = 1 | Z = z)}{P(E = e | Z = z)}.
\]

The last term in the right-hand side is the SMR weights given in the text. When the unexposed group is the target population, \( P(Z = z | \text{target population}) = P(Z = z | E = 0) \) and

\[
P(D = 1 | \text{set } E = e, E = 0) = \frac{1}{P(E = 0)} \sum_z P(D = 1, E = e, Z = z) \frac{P(E = 0 | Z = z)}{P(E = e | Z = z)}.
\]

**Appendix 2**

In this appendix, we provide SAS code to obtain the nonparametric multivariate extension of SMR described in the text. The following code is organized as follows. First, we use Proc Logistic to fit the logistic model in the estimation of propensity scores. Second, we use a SAS data step to calculate the proposed SMR weights for each subject from the estimated propensity scores of previous logistic model. Last, we use Proc Genmod to fit the weighted Poisson regression model that estimates the causal rate ratio.

The data file (SMR1) contains 1 record per woman. In the following code, the variable “ID” is the patient identification number and the variable “TAM” is a binary variable indicating whether a patient was exposed (TAM = 1) or unexposed (TAM = 0). The variables “AGE,” “STAGE,” “LYM,” and “MENO” are age at surgery, stage of breast cancer, lymph node metastasis, and menopausal status, respectively. The variable “RECUR” is the binary outcome of recurrence. The variable “LOGPT” is the logarithm of observed person-times for each subject.

```sas
/* Estimation of Propensity Scores */
proc logistic data=SMR1;
    model TAM=AGE STAGE LYM MENO;
    output out=PRED p=P1;
run;

/* Calculation of the SMR weights */
data SMR2;
    set PRED;
    if TAM=1 then WEIGHT=1;
    else WEIGHT=ODDS;
run;

/* Weighted analysis */
proc genmod data=SMR2;
    class ID;
    model RECUR=TAM
        /offset=LOGPT dist=poisson link=log;
    weight WEIGHT;
    repeated sub=ID / type=ind;
    estimate ‘Beta’ TAM 1 / exp;
run;
```

**REFERENCES**