
The original procedure

The spouses were grouped according to time (D) until – or since – entry into relationship. There were 16 groups with mean D varying from -12 years (before entry) to 39 (after entry). First we calculated polychoric spouse correlations for three variables (alcohol consumption, smoking, exercise) in each of the 16 groups, using Prelis, which also gives the asymptotic variances of the estimates. The correlation values were transformed to z-scores by Fisher’s formula, \( z(r) = 0.5 \times (\ln(1+r) - \ln(1-r)) \). We then generated a data file, using SPSS, in which we weighted each data observation (each z-transformed correlation) with the inverse of the variance of the estimates. For instance, the polychoric correlation for one group was 0.192, corresponding to \( z(r) = 0.194 \), with variance (squared standard error) of the estimate 0.00647. This standard error for a z-distributed variable corresponds to a sample with \( n=155 \), thus, this data point was weighted by the number 155.

Various functions, specified in the paper, were fitted to \( z(r) \), giving predicted spouse correlations as a function of D.

The error was made in the way we weighted the data points. Our weighing was equivalent with reproducing a record with the same value for the observed \( z(r) \) (in the example above, 155 times) which gives a falsely high precision.

How we corrected the method

Having consulted expert statisticians, we added an error term to \( z(r) \) (generated by the computer program), such that rather than generating a constant value (in the example 0.194 in all the 155 records), we generated a randomly fluctuating normally distributed variable with SD=1 (and, in the example, mean=0.194 for the group of the 155 records). Data were generated (simulated) in the same way for all sixteen observed correlations.
Unlike SPSS, the computer program R offers a nonlinear regression program with an option for weighting each observation. We have rerun the analysis in R, using this option. The results are almost identical to the new results from SPSS. The old and the new results (from SPSS) are shown below.

**Original and new, corrected results**

The parameter estimates from the old and the new analyses are practically identical, but the new confidence intervals are much wider. Not all parameter estimates remain significant. In order to demonstrate significant convergence prior to marriage, we specified reduced models with only one parameter for convergence before marriage - a linear regression coefficient. Results from these simpler models show significant increases (convergence) before marriage for all the variables. The graphical figures for the reduced models are not shown. They depart little from the original figures, except for showing straight instead of curved lines before marriage.

**Alcohol consumption**

Earlier published parameter estimates (with 95% CIs below):

\[
F1: y = 0.287 \times \exp(0.206 \times D_{neg}) + 0.0106 \times \ln(1 + 100 \times D_{pos}) + 0.239
\]

\[
\begin{align*}
&\text{(0.272--0.302)} \quad \text{(0.179--0.234)} \quad \text{(0.0093--0.0118)} \quad \text{(-15--216)} \quad \text{(0.223--0.254)}
\end{align*}
\]

New parameter estimates from the same model (with corrected 95% CIs below):

\[
F1: y = 0.289 \times \exp(0.205 \times D_{neg}) + 0.0105 \times \ln(1 + 100 \times D_{pos}) + 0.237
\]

\[
\begin{align*}
&\text{(-0.029--0.060)} \quad \text{(-0.379--0.788)} \quad \text{(-0.0161--0.0371)} \quad \text{(-2378--2577)} \quad \text{(-0.100--0.575)}
\end{align*}
\]

Parameter estimates from the reduced model with only a linear convergence before marriage:

\[
\begin{align*}
y = 0.023 \times D_{neg} + 0.0151 \times \ln(1 + 100 \times D_{pos}) + 0.490
\end{align*}
\]

\[
\begin{align*}
&\text{(0.006--0.041)} \quad \text{(-0.0115--0.0418)} \quad \text{(-1445--1645)} \quad \text{(0.377--0.604)}
\end{align*}
\]

The results from this model showed a significant linear increase in similarity during the 12 years before entry into marriage/cohabitation (t=2.59, p=0.0048).

**Smoking**

Published parameter estimates (with 95% CIs below):

\[
F3: y = 0.506 \times \exp(0.049 \times D_{neg}) - 0.000188 \times (D_{pos} - 0.0)^2 + 0.120
\]

\[
\begin{align*}
&\text{(0.417--0.595)} \quad \text{(0.038--0.060)} \quad \text{(0.000184--0.000192)} \quad \text{(-0.5--0.5)} \quad \text{(0.030--0.209)}
\end{align*}
\]
New parameter estimates from the same model (with corrected 95% CIs below):

\[ F_3 : y = 0.508 \times \exp(-0.049 \times D_{neg}) - 0.000188 \times (D_{pos} - 0.0)^2 + 0.117 \]

Results from a new, reduced model with a linear convergence before marriage:

\[ y = 0.020 \times D_{neg} - 0.000187 \times (D_{pos} - 0.0)^2 + 0.625 \]

The results from this model showed a significant linear increase in similarity in the time before entry into marriage/cohabitation (t=4.29, p<0.0001).

**Exercise**

Earlier published parameter estimates (with 95% CIs below):

\[ F_3 : y = 0.093 \times \exp(-0.274 \times D_{neg}) + 0.000204 \times (D_{pos} - 11.67)^2 + 0.189 \]

New parameter estimates (with 95% CIs below):

\[ F_3 : y = 0.093 \times \exp(-0.274 \times D_{neg}) + 0.000203 \times (D_{pos} - 11.67)^2 + 0.188 \]

A simplified model with a linear convergence before marriage:

\[ y = 0.093 \times \exp(-0.274 \times D_{neg}) + 0.000203 \times (D_{pos} - 11.67)^2 + 0.188 \]

There was a significant linear increase in similarity in the time before entry into marriage/cohabitation (t=1.74, p=0.041).

We are grateful to Dr. Håkon Gjessing and Dr. Bo Engdahl for valuable advice in correcting the results.