

**Step 1: Defining populations from which to draw data.**

**Population 1 (see Figure 2):**

The predictor and the health measure were associated at baseline due to factors with time-limited effects. These caused RTM in the health measure and were also related to the predictor. Hence, RTM in the health measure was related to the predictor. To avoid attributing RTM in the health measure to the predictor, the statistical model used for analyzing data from this population should account for RTM, as MR does.

**Population 2 (see Figure 3):**

The predictor and the health measure were associated at baseline due to factors with enduring effects. These did not cause RTM in the health measure. Hence, the predictor was not related to RTM in the health measure, and the statistical model used should therefore not assume RTM, as change score analysis does not.

**Step 2: Defining associations between attrition and the variables in the populations.**

Associations between liability of dropping out of the study and the other variables were modelled into Population 1 and Population 2. A latent variable defined liability of dropping out, and this variable was related to the study variables to varying degrees ( $b_a$  from .00 to .30). This allowed examining the effect of attrition with different degrees of dependency between attrition and the study variables.

**Step 3: Analyzing data drawn from the populations.**

500 samples were drawn from each of the defined populations. Each sample consisted of 1000 observations. To study bias related to selective attrition, some of the observations (30%, 50%, and 70%, respectively) were excluded from the analyses. MR was used for data from population 1 and change score analysis for data from population 2, to examine bias only related to selective attrition.

To study the combined effect of selective attrition and inappropriate modelling of RTM, the data were analyzed again, but this time data from population 1 were analyzed with change score analysis and data from population 2 with MR.

To decide the degree to which results were biased, they were compared to the population value of  $b_{pred} = .10$ .