Additional file 13. Five-fold cross-validation for the final models of voluntary pseudo-panel datasets.

|                       | Voluntary GZ Subset | Voluntary CD Subset |
|-----------------------|---------------------|---------------------|
| RMSE1                 | 1.1226              | 0.6458              |
| RMSE2                 | 0.8873              | 0.6756              |
| RMSE3                 | 1.0672              | 0.7114              |
| RMSE4                 | 1.0631              | 0.5327              |
| RMSE5                 | 0.8710              | 0.6045              |
| Average RMSE<sub>(cross-validation)</sub><sup>a</sup> | 1.0313              | 0.6370              |
| RMSE<sub>(full sample)</sub><sup>b</sup>         | 0.9856              | 0.6329              |
| ∆RMSE<sup>c</sup>     | 0.0457              | 0.0041              |
| % change in RMSE<sup>c</sup> | 4.6368              | 0.6478              |

<sup>a</sup>Average RMSE<sub>(cross-validation)</sub> = SQRT((RMSE<sub>1</sub><sup>2</sup> + RMSE<sub>2</sub><sup>2</sup> + ... + RMSE<sub>5</sub><sup>2</sup>)/5).

<sup>b</sup>∆RMSE = Average RMSE<sub>(cross-validation)</sub> - RMSE<sub>(full sample)</sub>.

<sup>c</sup>% change in RMSE = ∆RMSE×100/ RMSE<sub>(full sample)</sub>.