

International Journal of Population Data Science

Journal Website: www.ijpds.org



A Bayesian Record Linkage Approach that Adjusts for Variables in One File

Gauri Kamat¹, Mingyang Shan², and Roei Gutman¹

¹Brown University

²Eli Lilly and Company

In many healthcare and social science applications, information about units is dispersed across multiple data files. Linking records across files is necessary to estimate associations between variables exclusive to each of the files. Common record linkage algorithms only rely on similarities between linking variables common to all the files. Moreover, analysis of linked files often ignores errors that may arise from incorrect or missed links. Bayesian record linking methods allow for natural propagation of linkage errors, by jointly sampling the linkage structure and the model parameters. We extend an existing Bayesian record linkage approach to integrate associations between variables exclusive to each file being linked. We show analytically, and using simulations, that the proposed method improves the linkage process, and results in accurate statistical inferences. We apply the proposed method to link Meals on Wheels (MOW) recipients to Medicare Enrollment records, and examine the relationship between activities of daily living and healthcare utilization among MOW recipients.



Temporary page!

$\LaTeX$ was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, because $\LaTeX$ now knows how many pages to expect for this document.