**READ ME**

These materials relate to the following publication:

Woodhead, C., Martin, P., Osborn, D., Barratt, H., & Raine, R. (2021). Health system influences on potentially avoidable hospital admissions by secondary mental health service use: A national ecological study. Journal of Health Services Research & Policy, 13558196211036739 . <https://journals.sagepub.com/doi/full/10.1177/13558196211036739>

**Contact** about deposited data and code: Dr Peter Martin, Lecturer in Applied Statistics, University College London, peter.martin@ucl.ac.uk

The materials are designed to enable independent researchers to reproduce the analyses presented in the section “Predictors of variation in CCG-level PAAs”. We are not permitted to share HES data and thus have not included data or code to reproduce the calculation of Potentially Avoidable Admission (PAA) rates. Only the prediction of PAA rates by CCG characteristics is covered in the code presented.

**Data**

* The main data set to be read into R is: “MH\_Avoidable.csv”
* Explanation of variables in: “Avoidable Admissions - List of variables.xlsx”
* Lists of CCG identifiers that completeness thresholds for diagnostic information in the MHSDS data set (for sensitivity analyses) are:
	+ “CCGs with fewer than 50pct missing.csv”
	+ “CCGs with fewer than 70pct missing.csv”

**R code**

To run R code, first open the R Studio project “Potentially avoidable admissions - data and code.Rproj”.

* Code for implementing MIRL to investigate predictors of PAA rates:
	+ in secondary mental health service users: “MIRL stages 2,3,4 - MHSDS patients - physical admissions.R”
	+ in the comparator group: “MIRL stages 2,3,4 - No MHSDS - physical admissions.R”
* code for sensitivity analyses:
	+ “Sensitivity analysis.R”
	+ “Sensitivity analysis 70pct missing.R”
* code for fractional polynomials: “MH avoidable – fractional polynomials.R”

All other R code files contained within the project are called from within these five files via the source command.

**Random seed compatibility**

The R code for this project was written in R version 3.5.0, and the results were obtained using this version. It appears that from R version 3.6.0, random numbers generated via the set.seed() command are not necessarily compatible with previous versions. It is thus possible that future analysts will obtain slightly different results than those published in the paper, even if using the same seed numbers. See: <https://stackoverflow.com/questions/47199415/is-set-seed-consistent-over-different-versions-of-r-and-ubuntu> for more information.

Peter Martin, October 2021