# AMR-UCR Code Manual

Use and citation

For general usage and citation see the README file. This manual contains brief explanations of the documents and specific notes on inputs/outputs, and information on input file code (in the absence of a separate data dictionary).

Data Extraction File (cost\_per\_case/inputs/DET.xlsx)

* For Region;
  + Europe; may include EU and/or wider Europe (all grouped under Europe as a region) in this version of the repository
* For microbe;
  + Efa; included Enterococcus not necessarily just Eneterococcus faecium
  + If a study grouped E. coli and Klebsiella together this was put down as Enterobacteriaceae
* For cost year;
  + If a specific year wasn’t specified then took the latest study period year for costs
* For syndrome;
  + BSI; bloodstream infection
  + B-J-SSTI; bone, joint and skin and/or soft tissue infections
  + COL; exposures classified as colonized
  + COL/INF; exposures included colonized and infected
  + IAI; intra-abdominal infection
  + INF; exposures classified as infected
  + RTI; respiratory infection
  + UTI; urinary tract infection
* For LOS/cost extractions;
  + Extracted post-infection hospital if available, if not total hospital (not just pre-infection LOS), and if not hospital then ICU.
* For measure;
  + “excess” = “excess mean”, unless specified otherwise (e.g. “excess\_median” means excess meadian)
* For the flags, which were used to do checks of data extractions: Note “FALSE” values meant no further checks required based on the flag.
  + FLAG\_COMPLETE 🡪 that the descriptor columns were filled
  + FLAG\_MONEY 🡪 to flag that if it was a LOS estimate row the currency columns weren’t filled in Or if it was a cost estimate that it was filled in, note this was done by hand subsetting los=1 and los=0 and applying different formulae to each (so you can’t drag and drop across the whole data set the formulae).
  + FLAG\_EXPOSED<NON 🡪 to flag if nonexposed values were greater than exposed, this didn’t mean it was wrong, but may warrant another check of the source to make sure the correct values were assigned to the correct groups.
  + Additionally unique\_id\_4both was added in by hand throughout the extraction process, and subsequently checked by hand, due to removals of rows in the data checking and cleaning process there may be gaps in numbers. For example…

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| first author(s) | study year | both\_los\_cost | | unique\_id | |
|  |  |  |  | |
| Kim | 2014 | 1 | 4 | |
| Kim | 2014 | 1 | 4 | |
| Klein | 2019 | 1 | 11 | |
| Klein | 2019 | 1 | 11 | |

* + This is used in the code where we want to compare the cost difference of just costing length of stay vs using their own cost differences in studies which had both.