**Chronic Myeloid Leukemia: Two Mysteries**

Tomas Radivoyevitch, Trine N. Jorgensen, Daniel J. Lindner,

Jaroslaw P. Maciejewski, Yasushi Miyazaki and Robert Peter Gale

**Software and Data Installation Instructions**

To reproduce our results software and data must be installed. First install [R](https://cran.case.edu/) and [RStudio](https://www.rstudio.com/). Then install SEERaBomb by typing install.packages("SEERaBomb") at the R prompt, or, to get the most recent version via GitHub, by using devtools::install\_github("radivot/SEERaBomb",subdir="SEERaBomb").

Next get the [SEER ASCII *treatment* data](https://seer.cancer.gov/data/treatment.html) by signing an additional radiation and chemotherapy data use agreement and unzip it to ~/data/SEER where ~ is your home directory. Now generate SEER binary data via

library(SEERaBomb) #loads installed package SEERaBomb into memory

(df=getFields()) #gets SEER fields into a data frame

(rdf=pickFields(df))#picks a subset of SEER fields and defines their types

mkSEER(rdf) #makes merged data file ~/data/SEER/mrgd/cancDef.Rdata

and check the installation using

load("~/data/SEER/mrgd/cancDef.RData")#loads data.frame canc into memory

head(canc,3) #returns top 3 rows of canc

Obtain the data [“*The incidence of leukemia, lymphoma and multiple myeloma among atomic bomb survivors: 1950-2001*”](https://www.rerf.or.jp/en/library/data-en/lsshempy/) and [“*Life Span Study Cancer Incidence Data, 1958-1998*”](https://www.rerf.or.jp/en/library/data-en/lssinc07/) from the Radiation Effects Research Foundation and place lsshempy.csv and lssinc07.csv in ~/data/abomb and run

mkAbomb()#converts files in ~/data/abomb into ~/data/abomb/abomb.RData

Install other R packages using install.packages(c("bbmle","ggsci","tidyverse"))

Running the R scripts in <https://github.com/radivot/SEERaBomb/tree/master/SEERaBomb/inst/docs/papers/CMLmysteries>

should now regenerate Figure 1 and Table 1 and other calculations in the text.