



City Research Online

City St George's, University of London

Citation: Reyes-Aldasoro, C. C. (2017). The proportion of cancer related entries in PubMed has increased considerably. Paper presented at the 2017 NCRI Conference, 5-8 Nov 2017, Liverpool, UK.

This is the supplemental version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/18496/>

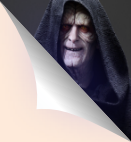
Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).



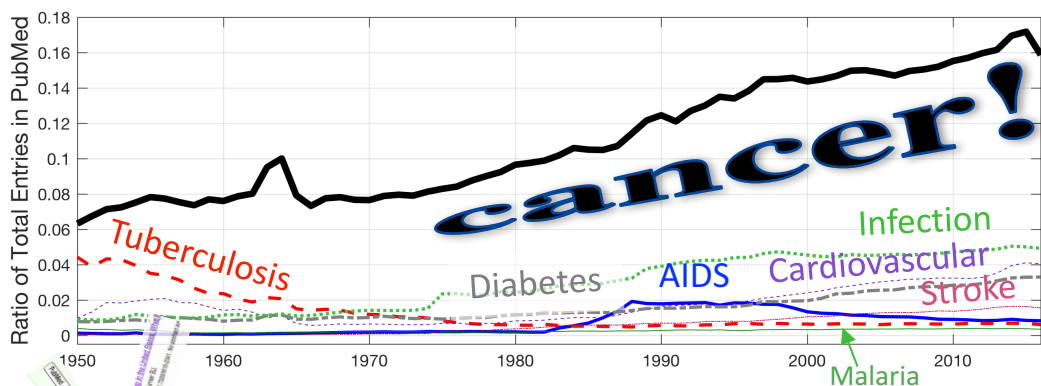
The proportion of cancer-related entries in *PubMed* has increased considerably; is cancer truly *The Emperor of All Maladies*?

Constantino Carlos Reyes-Aldasoro

School of Mathematics, Computer Science and Engineering, City, University of London, EC1V 0HB



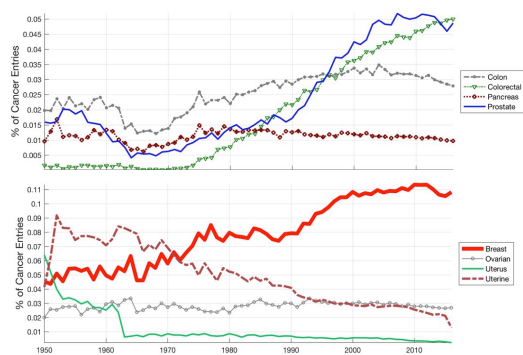
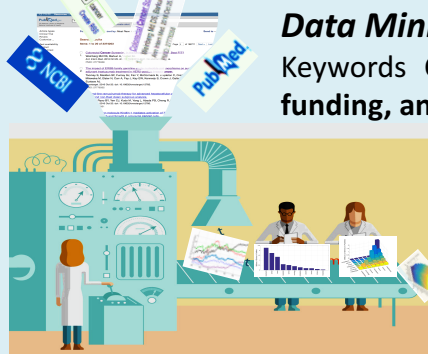
The public database of biomedical literature *PubMed* was data-mined using queries with combinations of keywords and year restrictions. It was found that the proportion of *Cancer-related* entries per year in PubMed has risen from around **6% in 1950** to more than **16% in 2016**. This increase is not shared by other conditions such as *AIDS*, *Malaria*, *Tuberculosis*, *Diabetes*, *Cardiovascular*, *Stroke* and *Infection* some of which have, on the contrary, decreased as a proportion of the total entries per year.



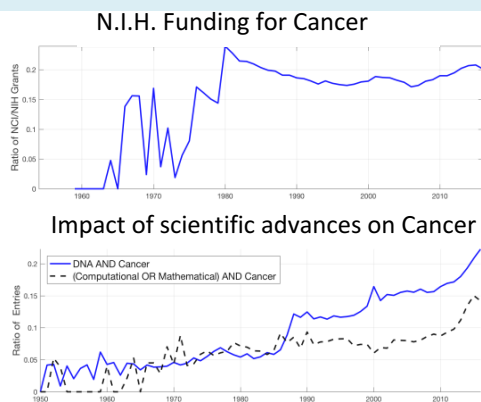
Data Mining Approach:

Keywords **Cancer terminology, organ-related, incidence, funding, and relationship with DNA, Computing**

and Mathematics. Interestingly, the proportion of Cancer-related entries that contain “DNA”, “Computational OR Mathematical” have increased, which suggests that the impact of these scientific advances on Cancer has been stronger than in other conditions.



Organ-specific research showed mixed trends



The sharp increase of Cancer Research as testified by the number of entries in PubMed may be due to the strong impact of the scientific advances in the areas of **Genetics, Computing and Mathematics**, which have had a *stronger influence* in Cancer than other areas like cardiovascular disease. It is important to highlight that the results obtained with a *data mining* approach and thus are limited to the presence or absence of the keywords on a single, yet extensive, database.

Reyes-Aldasoro CC (2017) The proportion of cancer-related entries in PubMed has increased considerably; is cancer truly “The Emperor of All Maladies”? *PLOS ONE* 12(3): e0173671. <https://doi.org/10.1371/journal.pone.0173671>



@C_ReyesAldasoro #NCRI2017