

## SUPPLEMENTARY FILE

# **Female and Male United States Olympic Athletes Live 5 years longer than their general population counterparts: A study of 8,124 former US Olympians.**

J Antero *Ph.D.*<sup>1</sup>, H Tanaka *Ph.D.*<sup>2</sup>, Q De Laroche Lambert *M.S.*<sup>1</sup>,  
M Pohar-Perme *Ph.D.*<sup>3</sup>, and JF Toussaint *M.D. Ph.D.*<sup>1,4,5</sup>

<sup>1</sup>Institute for Research in Medicine and Epidemiology of Sports (IRMES, EA 7329), The French National Sports Institute (INSEP), Paris, France

<sup>2</sup>The University of Texas at Austin, Austin, TX USA

<sup>3</sup>Institute of Biostatistics and Medical Informatics, Medical faculty, The University of Ljubljana, Ljubljana, Slovenia

<sup>4</sup>Paris Descartes University, EA 7329, Sorbonne Paris Cité, Paris, France

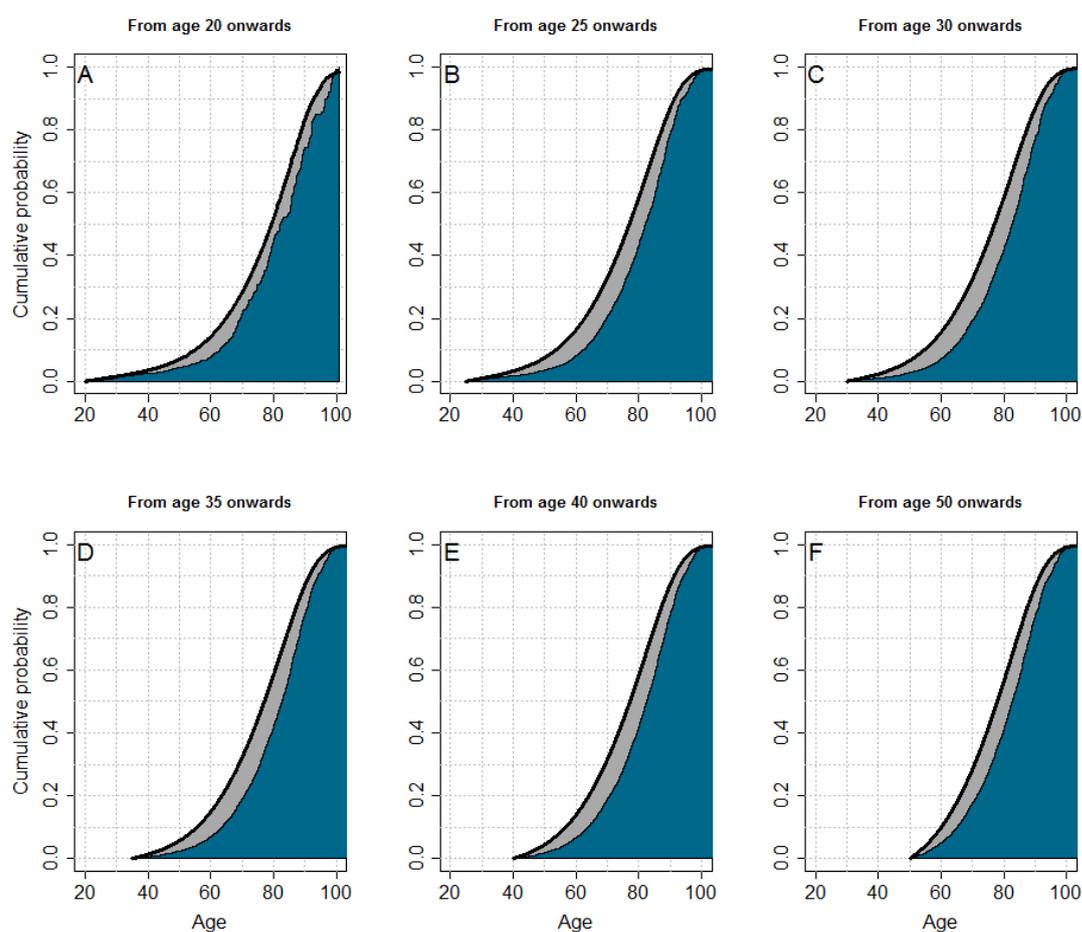
<sup>5</sup>CIMS, Hôtel-Dieu, AP-HP, Paris, France

**TABLE OF CONTENTS**

The effect of age at the cohort entry.....	3
Sensitivity analyses of the imputed data.....	5
Premature deaths in period 3 (post WWII, 1948-1964) .....	7

### The effect of the age at the cohort' entry

We assumed that at a certain age, all individuals have the same risk regardless of the age they competed in their first Olympiad. To test the effect of age at the cohort entry we analyzed the cumulative probability of death considering only individuals who had already been at Olympics by the age of 20, 25, 30, 35, 40 and 50 years (Figure S1).



**Figure S1:** Cumulative probability of death conditional to the age of first Olympic participation. A) From age 20 onwards. B) From age 25 onwards. C) From age 30 onwards. D) From age 35 onwards. E) From age 40 onwards. F) From age 50 onwards.

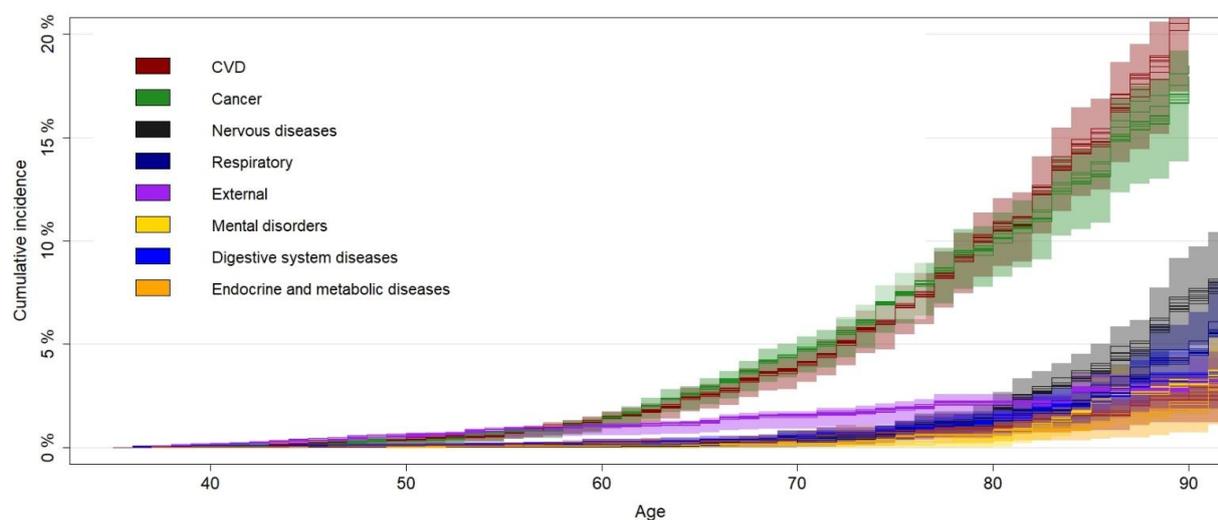
The cumulative probability of death is similar for every group of age, with the exception of the youngest group (which is based on the smallest sample). There is little gain in the number of years saved before turning 40, since mortality in the general population is also low.

With the age of the athletes as the time scale of interest, the cause of death analyses were performed conditional on athletes surviving to a certain age and excluding those who competed in the OG at a later age. The age of 35 was chosen, as this age represents the best concession between a loss of information (for athletes competing at older ages) and the longest follow-up.

### Sensitivity analyses of the imputed data

The cumulative incidence of the different causes are represented after imputation in figure S2. All the 8 data sets built with imputed data for missing causes of deaths show similar cumulative incidences: all cumulative incidence curves are comprised in the 95% CI of the original data. This illustrates a high concordance and the robustness of the chosen imputation model.

In addition, the missing causes of death are not related with the causes itself. For the deaths occurring before 1979 (42%), data are missing due to the lack of official sources to confirm death certificates. For the missing identification of cause of deaths occurring after this period (58%), data are missing completely at random, mostly related to the communality of the athletes' names. We observed that the more unusual their names were, the easier their



identification in the NDI database.

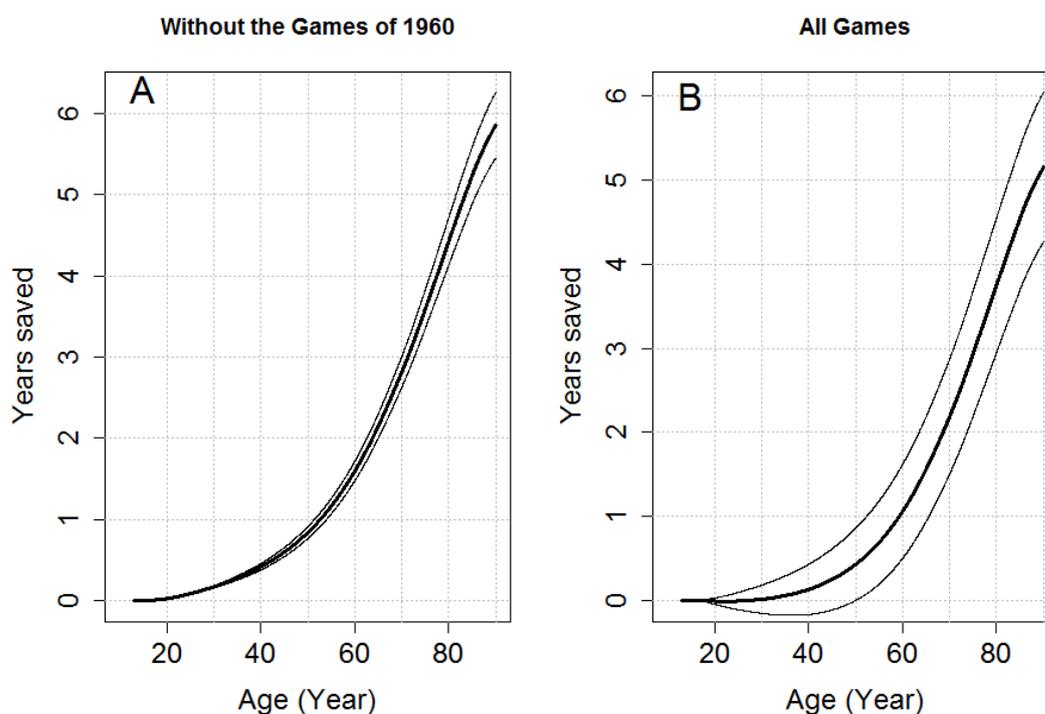
**Figure S2:** The cumulative incidence of causes of deaths for each of the eight different datasets created after imputation (thin lines). The 95% CI (shaded areas) is based on the data before imputation.



### Premature deaths in the period 3 (post WWII, 1948-1964)

We observed an unusual number of deaths among young Olympians during period 3, occurring simultaneously and related to external causes. All these deaths concerned Olympians who participated to the 1960's Games. The single reason for these deaths relates to a plane crash, killing the entire US figure skating team one year after their participation in the Games, as confirmed through public sources <sup>1</sup>.

The effect of such an incident in the years-saved analysis is illustrated in Figure S3.



**Figure S3:** Years-saved among US Olympians. A) Excluding all the athletes ( $n = 270$ ) participating in the 1960 Olympic Games. B) All US Olympians ( $n = 8124$ ).

We observed that the 95% CI is much narrower when excluding the 1960 Games, resulting in a total of 5.8 years-saved (5.5 to 6.3). This single accident affects the analyses

because the athletes were very young (under 20 years old) and mostly women, for which the mortality rates in the general population are extremely low.

Acknowledging the atypicality of these deaths, we still could not exclude them as they were associated with the status of Olympian (they all were heading to a competition). Yet, the analysis of the causes was not affected by this incident, since it starts at 35 years old.

**References:**

1. In 1961 a Plane Crash Killed the Entire U.S. Figure Skating Team. *Time*. <http://time.com/5172201/1961-plane-crash-us-figure-skating/>. Accessed November 20, 2018.