Anthropometric Data Analytics: Portuguese case-study

António Barata[1], Francisco M. Couto[1], Lucília Carvalho[2]

[1]LaSIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2]Hospital de Egas Moniz, Centro Hospitalar de Lisboa Ocidental, Portugal

Large amounts of information are systematically generated throughout the course of scientific research and progress. In our case, over 10 years of data were collected by means of measuring distances and weights of numerous anthropometric features and organs, respectively, through various fetal-autopsy procedures. With such a substantial sample size, it is possible to produce a statistically-inferred table with the expected values, following a pre-defined confidence interval, of each accounted variable for each week of gestational age. To this day, Portuguese medical professionals have no access to such a unifying construct, having to rely on foreign lists and approximations to assess the developmental stage of each individual. To address this issue, we developed a system that reads from autopsy report files in pdf format and retrieves the values of each accounted variable, thus, alongside manual inputting, simple python scripts are used to create a straightforward SQLite database containing information about every individual. This database enables us to compute, for each sampled variable per gestation period, the corresponding population mean and standard deviation values, following a Student's t-distribution and chi-squared distribution, respectively, due to the absence of known population parameters. This organized information scheme is to replace the currently employed North-American table, enabling a simpler and more accurate way to process and retrieve information by fetopathology and developmental pathology professionals in Portugal. As further work is developed, we aspire to create an online query system which, supplied with certain inputs such as values for the variables collected and studied in this work, ensures a list of links to the relevant bibliography for that specific query with the possibility of user-feedback being used as a machine-learning mechanism in order to improve this system; in layman's terms, to create a tool meant to not only diagnose certain pathological conditions based entirely on anthropometric measures but also to provide relevant information associated with the stipulated medical condition assessed previously, trimmed by the users themselves.

Preference for presentation: Poster
Location: University of Minho
Author for Correspondence: apbarata@gmail.com