The use of Big Data in the understanding the epidemic of chronic kidney disease in Central America.

Recent developments about Big Data provide new challenges and opportunities for research. The ability to collect and analyze large volumes of data brings about a new perspective in research. Organisations and researchers in different fields are now able to extract insight and patterns from large datasets because it is now easier to collect large datasets either as part of their research or use data made available from previous research. Moreover, the availability of new analytical tools and techniques makes it possible to combine several structured and unstructured datasets to create new knowledge. The use of Big Data and analytics will be applied in this study for profiling and understanding the populations that are at risk, to predict those that might be at risk in the future and to determine the most appropriate solution for the different categories within the target areas.

Chronic kidney disease (CKD) not related to traditional causes (diabetes and hypertension) has reached epidemic proportions in Central America; affects mainly young male agricultural workers but also, to a lesser extent, women and non-agricultural workers living in farming communities of Central America. El Salvador and Nicaragua possess the highest mortality rates due to CKD and data suggest this will increase in the following years. The two main hypotheses that explain this epidemic are: 1) indiscriminate use of agrochemical products, and 2) heat stress along with dehydration. Both are related to a lack of a regulatory system to control agrochemical use and the poor compliance with rules and standards to protect the labor force’s health. El Salvador has leaded the research in the region of Central America showing several results in the pasts few years, but the causes of the epidemic are not completely understood.

Big Data analysis can help to understand the behavior of this disease by comparing epidemiological data between different countries (regardless their development status) creating a comparative analysis of several trends. Several layers of data ranging from demographic, economic, infrastructure, health-related, and clinical data can merge together to create a panorama of the disease in different countries-leading to a better planning of resources. The US federal government and public stakeholders have been opening their health-databases including data from clinical trials and information on patients covered by insurance programs and are available to the public and researchers.

Reference:

Ordunez P, Saenz C, Martinez R, Chapman E, Reveiz L, Becerra F. [The epidemic of chronic kidney disease in Central America.](http://www.thelancet.com/journals/langlo/article/PIIS2214-109X(14)70217-7/fulltext) The Lancet Global Health 2014. Early Online Publication 25 June 2014 doi:10.1016/S2214-109X(14)70217-7

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. 2011. Big data: The next frontier for innovation, competition and productivity. Available at: <http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation>