Dynamical augment in mining healthcare datasets

John Heironimus[1] and Emma Heironimus[2]
[1]) O H P L Q J 6 F L H Q W L À F 8 6 $
[2]University of Louisville, Kentucky

Statistical processes commonly applied to healthcare datasets can overlook important dynamical relationships. However, understanding dynamicals o en involves complex and expensive modelling. We demonstrate a simple approach to "dynamical data mining" using Fourier transform and the Kuramoto model. We compare results from this approach to cross correlations using CDC's ICD-113 data (1999-2016). We nd that the ICD-113 is a synchronous anti-phase nodal system much like the spontaneous synchrony of pendulum clocks observed by Christiaan Huygens in 1665 (see image). It is a sta of order in which ICDs are related phasallyabout common frequencies. We nd that about 20% of these phasalrelationships are novel to the statistical approach and corroborated by research ndings. Many of these novel insights appear to be comple time series relationships suggesting stronger linkages between initial conditions/care and nal causes of death than statistic would typically reveal. ICD-113's state of order remains stable even as the mix of ICDs changes. is suggests that declining US healthcare productivityis perhaps more attributable to complex dynamical relationships within the system than to rising obesity or opioid abuse rates as o en hypothesized. Given the system's stablestate of order, we can simulate the impact changing incidence of any one condition on all others without consideration of causality. As a synchronous system can re ect the in uence of external factors, we also consider the possible in uence of solar cycles. Using the SILSO dataset, we nd consistency with ICD-113 (see image). Each ICD's phasalrelationship to individual solar cycles allows us to infer its most likely time series relationship and possibly forecast future incidence on the same basis. We o er several detailed examples all points in our paper and conclude that simple "dynamical mining" o ers an important augment to statistical processing of healthcare datasets.

Biography

Notes: