Comment

Facts and figures

Gregory A Petsko

Address: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110, USA.
Email: petsko@brandeis.edu

Published: 28 July 2006

Genome ***Biology*** 2006, **7:**111 (doi:10.1186/gb-2006-7-7-111)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2006/7/7/111

I spend far too much time flipping through a slim volume that arrives in my mailbox at the start of every summer. No, it isn't a travel brochure for island get-aways. It's the Pocket World in Figures, published by *The Economist* magazine. (I realize I am committing something of a *faux pas* by referring to *The Economist* as a magazine. With a haughtiness they perhaps have earned, since they've been in business since September 1843, they always refer to themselves as "this newspaper". It's sort of like David Beckham referring to himself in the third person, I guess.) I've been a subscriber for over 30 years, and along with the only decent news coverage of the world outside the United States, one of the other benefits I get is this yearly little book.

You can learn all kinds of amazing things just from opening it at random. For example, the United States has the second highest divorce rate in the world: 4.8 divorces per 1,000 population (the UK is 14th, at 3.0; Canada is 29th at 2.3). The country with the highest rate is Aruba (5.3). Aruba? What on earth is it about Aruba that puts such a strain on marriages? It can't be the tourists: no other Caribbean nation makes the top 40. Certainly can't be the weather. Aruba touts itself as a very friendly place; maybe the people there are too friendly. Anyway, whatever the reason, there's something about Aruba that's apparently incompatible with marital longevity. The country with the lowest rate is Colombia, with only 0.2 divorces per 1,000 people. The fact that it's a Catholic country is certainly part of the reason, but so might be the fact that, owing to a rather serious crime problem, a large segment of the population is armed to the teeth. You might think twice about divorcing your spouse if you know that he - or she - regularly packs some serious firepower.

Particularly interesting are the statistics about health and disease. Swaziland has the highest death rate, 31.2 per 1,000 people. Places 2 through 18 on the list are also held by African countries where AIDS, malaria and/or tribal wars are endemic. The lowest death rate, 1.3 per 1,000, is in the United Arab Emirates. Places 2 through 10 from the lowest death rate are occupied by other oil-rich, Middle-Eastern countries, suggesting a correlation between wealth and health. But it can't be that: Mexico is 13th and the Philippines is 16th from the top. Neither the US nor the UK even cracks the top 25. The disease figures are equally fascinating. The countries with the highest death rates from breast cancer are all Northern European. The diets and lifestyles are so different, it makes one wonder if there's a gene in there somewhere. Hungary has a lung cancer death rate almost 20% higher than any other country on earth. And the US spends more on health care, as a percentage of Gross Domestic Product, than any other country by a large margin, yet it is tied with Portugal for 40th on the table of life-expectancy. Japan, where people live on average 5 years longer, is not even in the top 30 in health care spending. In fact, a plot of health care expenditure against either life expectancy or death rate is a scatter plot, with no apparent correlation whatsoever.

We could argue forever about what some of these facts mean, but I doubt there is a single scientist who wouldn't find the data interesting. Most of us become scientists because we are fascinated by the causes of things, and the causes of death and disease affect us all. Deducing these causes from a morass of figures is the work of one of my favorite sciences, and one that I think should be the closest ally of genomics: epidemiology. Epidemiology is the study of the factors involved in the health and illness of individuals and populations. It's the cornerstone of both public health and preventative medicine. Epidemiologists try to establish statistical relationships between disease agents, both infectious and non-infectious, and human illness: for example, they were the first to establish a link between lung cancer and cigarette smoking. Although human beings have probably speculated about such connections since the Stone Age, the science of epidemiology is only about 150 years old. It begins, as so much of our modern world does, with the Victorians.

The summer of 1854 saw the worst cholera epidemic in the history of London. Over 10,000 people died. It flared up, seemingly at random, in various spots throughout the crowded, unsanitary city. As the summer drew to a close, Soho, which had been spared up to that point, was suddenly hit hardest of all. The majority of those who died became ill on the night of 31 August and were dead within days. It is said that 75% of the population of the city left town that week out of fear of becoming ill. One who didn't leave, however, was London's most famous physician, Dr John Snow.

John Snow was born in 1813, the eldest son of a Yorkshire farmer. Apprenticed to a surgeon at 14, he graduated from the University of London with his medical doctorate on 20 December, 1844. He had long been fascinated by the mode of transmission of cholera, and as a young surgeon had actually written a paper, which was largely ignored, proposing that contaminated water supply might be the mechanism. In 1846, he became interested in reports from America about the properties of ether as an anaesthetizing agent. (The history of anesthesia is one of the most remarkable of all medical stories, and one that I plan to write about in the near future.) Experimenting largely on himself, Snow developed improved methods of administering the drug and demonstrated its effectiveness in the dental out-patient clinic at St George's Hospital. You might think that, as the man who introduced the use of ether into English surgery, Snow would have had a vested interest in the continued use of this drug. But he had a true scientist's mentality, and when data began to reach him that chloroform had advantages, he unhesitatingly championed its use. Anesthesia was scorned for use in child-birth by the (entirely male) medical establishment of that time, not because of suspicion of its effectiveness, but because it was thought to be contrary to the biblical injunction to Eve, "in sorrow thou shalt bring forth children (Genesis 3:16)". All that changed, forever, on the 7 April, 1853, when Snow persuaded Queen Victoria to use chloroform anesthesia for the birth of Prince Leopold, her eighth child. This, then, was the doctor who turned his attention to the cholera epidemic raging through London at the end of August 1854.

Medical wisdom at that time held that cholera was spread by 'miasma in the atmosphere' (I don't know quite what that means, but it sure sounds bad, doesn't it?). Because of that belief, nothing was done to contain an epidemic when it broke out. Snow, however, had already formed the opinion that contaminated water might be responsible, and the Soho epidemic gave him his chance. At that time he was living on Frith Street, right in the heart of the district. He started to patrol the district. He interviewed the families of victims, and began to plot on a map of that part of London the deaths from cholera. He found that nearly all of the deaths had taken place within a short distance of a water pump on the corner of Broad and Cambridge Streets. The well was nine meters deep, but a sewer only seven meters below ground was just above it. Snow took his findings to the Board of Guardians of St James's Parish, where the pump was located, and persuaded them to remove the pump handle so no one could obtain water there. The cholera outbreak stopped almost immediately.

Although the Victorian medical establishment continued to express skepticism about Snow's theory of cholera transmission, this remarkable demonstration of the power of data analysis led eventually to improvements in London's sanitation, including the construction of a proper sewerage system in the 1880s. John Snow is widely credited with the invention of the science of epidemiology. The site of the Broad Street pump, now the corner of Broadwick and Lexington Streets in Soho, is the location of a pub, aptly called The John Snow. A water pump with no handle sits outside. It was erected there on 28 July, 1992. A picture of the good doctor is on the pub's signboard. There's a room at the back on the second floor housing an exhibit dedicated to his life and work; if you ask, you'll be allowed to go see it. It's a good pub, too - a favorite hang-out of mine when I'm in London. You just might see me there when you visit. Ironically, you would never see John Snow there: he was a teetotaler. There's a society dedicated to promote knowledge of the life of John Snow and of his works, to encourage communication and collaboration between specialists of the many disciplines that have benefited from Snow's work, and to ensure that John Snow's memory continues to be celebrated in the pub bearing his name. I'm a member. Since March 2001, the John Snow Society has been based in the Royal Institute of Public Health, Portland Place, London. In 2003, John Snow was voted the greatest doctor of all time in a poll of physicians by the journal *Hospital Doctor*. Hippocrates finished second.

The science of epidemiology that Snow created remains, in my view, one of the most fascinating, and important, of the medical sciences. As for why I think it ought to be the closest allay of genomics, consider the following connections. Both sciences are data rich. Both depend on finding unexpected connections within data. Both aim to get at the underlying causes of things. And where genomics has been sold - I would claim, oversold for the present - as a key to personalized medicine, epidemiology is where the first glimmerings that such medicine might even be possible came to light.

If we could collect and mine enough epidemiological data, it would tell us where to look for the genes most important to human health. For example, several human genes that, when overexpressed, are associated with various cancers have the property that their deletion or loss of function is associated with neurodegenerative conditions like Parkinson's Disease. (Though surprising, the finding makes sense because cancer involves cells living that ought to die because of genomic instability or damage, while neurodegeneration involves the death of cells that ought to live: they are two sides of what might be the same coin.) This connection was discovered by the fact that cancer researchers doing proteomics found upregulation of genes in tumors that neurologists had found using human genetics and the genome sequence to be mutated in rare, autosomal recessive, familial Parkinson's Disease. Yet epidemiologists had already suspected that there was such a connection, because Parkinson's patients tend to have lower than normal incidence of certain cancers. I think one of the best places to look for new cancer genes is in the genetic neurologic disorders, and if we want to determine the pathways to neurodegeneration, we might want to look closely at genes that are important for tumor survival.

Similarly, consider the enormous number (hundreds) of inborn errors of metabolism. Most are recessive, and in many cases the carriers are asymptomatic. Considering that many of these conditions have carrier frequencies of 1 in 200 or even greater, the number of carriers in the general population is quite high. An obvious question, given the importance of most of the affected proteins, is whether being a carrier puts the individual at risk for anything else? Epidemiological evidence is mounting that it does. Gaucher Disease, the most common genetic disorder among Ashkenazi Jews, has a carrier frequency in the general population that is estimated to be as large as 1 in 100. Gaucher carriers show no symptoms of Gaucher Disease, but epidemiological data indicate that they may have as much as a 10-fold higher risk of sporadic, idiopathic Parkinson's Disease than non-carriers. I suspect that haploinsufficiency of important metabolic enzymes may underlie many of the 'sporadic' diseases. The right marriage of genomic and epidemiological data should find these connections, and with them, uncover the pathways to disease.

We need to think as broadly as possible about questions like this. We need the barriers between disciplines to be as low as possible. We need data from sciences like genomics and epidemiology to be freely available in forms that can be understood widely and transferred from one discipline to another easily. It may be that we need genomics and epidemiology to become one science. If we had genome sequence information for whole populations coupled with epidemiologic information about health and disease and lifestyle, imagine what a new generation of John Snows could do with such data. And imagine what a book of figures that would make!