Social and environmental risk factors in the emergence of infectious diseases

Robin A Weiss¹ & Anthony J McMichael²

Fifty years ago, the age-old scourge of infectious disease was receding in the developed world in response to improved public health measures, while the advent of antibiotics, better vaccines, insecticides and improved surveillance held the promise of eradicating residual problems. By the late twentieth century, however, an increase in the emergence and re-emergence of infectious diseases was evident in many parts of the world. This upturn looms as the fourth major transition in human–microbe relationships since the advent of agriculture around 10,000 years ago. About 30 new diseases have been identified, including Legionnaires’ disease, human immunodeficiency virus (HIV)/acquired immune deficiency syndrome (AIDS), hepatitis C, bovine spongiform encephalopathy (BSE)/variant Creutzfeldt-Jakob disease (vCJD), Nipah virus, several viral hemorrhagic fevers and, most recently, severe acute respiratory syndrome (SARS) and avian influenza. The emergence of these diseases, and resurgence of old ones like tuberculosis and cholera, reflects various changes in human ecology: rural-to-urban migration resulting in high-density peri-urban slums; increasing long-distance mobility and trade; the social disruption of war and conflict; changes in personal behavior; and, increasingly, human-induced global changes, including widespread forest clearance and climate change. Political ignorance, denial and obduracy (as with HIV/AIDS) further compound the risks. The use and misuse of medical technology also pose risks, such as drug-resistant microbes and contaminated equipment or biological medicines. A better understanding of the evolving social dynamics of emerging infectious diseases ought to help us to anticipate and hopefully ameliorate current and future risks.

Popular writing on emerging infectious diseases resounds with dire warnings about the threat of modern ‘plagues’ and losing the ‘war against microbes.’ This adversarial language obscures the fact that most of the microbial world is either neutral toward, or supportive of, human well-being and survival. Indeed, we would not survive long without commensal microbes such as the beneficial strains of Escherichia coli in our gut. That aside, the study of emerging infections is more than a passing fad. The recent rate of identification of such infections, the impact of the SARS outbreak, the devastation caused by AIDS, and the ever-present threat of a new influenza pandemic indicate that we cannot control our disease destiny. Nor are emerging infections unique to humans; the Irish potato famine in 1845 and the English foot-and-mouth disease epidemic in 2001 underscore the consequences for human societies of disease emergence in crops and livestock.

Emerging infectious diseases in humans comprise the following: first, established diseases undergoing increased incidence or geographic spread, for example, Tuberculosis and Dengue fever; second, newly discovered infections causing known diseases, for example, hepatitis C and Helicobacter pylori; and third, newly emerged diseases, for example, HIV/AIDS and SARS.

This Perspective will discuss the human ecology of both the (apparently) new and re-emerging diseases.

The demography of infectious disease

Interest in infectious disease has itself recently re-emerged. In 1972, Burnet and White commented, “The most likely forecast about the future of infectious disease is that it will be very dull. There may be some wholly unexpected emergence of a new and dangerous infectious disease, but nothing of the sort has marked the past fifty years”¹. Today, we may criticize the short-sightedness of our mentors’ generation, yet in demographic terms they were essentially correct because the proportion of deaths from infectious disease has fallen throughout the twentieth century²,³ (Fig. 1).

Humankind currently faces neither apocalyptic extinction nor even a population reduction such as occurred in Europe during the Black Death of the fourteenth century. Rather, overpopulation in relation to environmental resources remains a more pressing problem in many developing countries, where poor economic and social conditions go hand-in-hand with infectious disease. In industrializing countries during the nineteenth century, a major reduction in enteric infections was achieved by separating drinking water from sewage—an environmental change that probably saved more lives than all the twentieth century vaccines and antibiotics together. Today, however, the growth of shanty towns without sanitation around the megalopolis cities of Asia, Africa and South America is recreating similar conditions, and in the past 40 years cholera has made a remarkable re-emergence through its longest ever (seventh) pandemic⁴.

In most countries, life expectancy has risen over the past 50 years⁵ (Fig. 2). The most important exception is those regions where HIV infection is rife. Moreover, during the past 15 years, falling living standards in some African countries and the breakdown of public health

¹Wohl Virion Centre, Division of Infection and Immunity, University College London, W1T 4JF, UK. ²National Centre for Epidemiology and Population Health, The Australian National University, Canberra, ACT 0200, Australia. Correspondence should be addressed to R.A.W. (r.weiss@ucl.ac.uk) or A.J.M. (tony.mcmichael@anu.edu.au).
infrastructure in ex-Soviet nations has aided the re-emergence of transmissible diseases like tuberculosis\(^4\,6\). Further, severe outbreaks such as the 1918–1919 influenza A pandemic temporarily reversed the decline of deaths caused by infectious disease. The 50 million estimated deaths from that pandemic\(^7\) represented about 2\% of the global population at that time, and is twice as many as the cumulative AIDS mortality of the past 20 years. The next influenza pandemic may be just around the corner, and may spread even faster\(^8\,9\,10\), if access to appropriate vaccines and drug treatment is not available\(^9\,10\).

For other newly emerging infections that make headlines, such as SARS, Ebola or vCJD, it is important to keep a sense of demographic proportion. Placing these emerging infections on a ‘Richter’ scale of human mortality (Box 1) shows that they elicit scarcely detectable minor tremors in numbers of fatalities — despite the fear they invoke. We do not know, however, which one might leap to the top of the scale like HIV has done; indeed, it may be a completely unknown agent, as the SARS coronavirus was two years ago. A major challenge is to predict which infection presages the next big quake, hopefully allowing preventive action.

Emerging infectious diseases, past and present

Like any other animal or plant species, humans have been prone to infection by pathogens throughout their evolutionary history. Such ancient infections by helminth and protozoan parasites, bacteria, fungi and viruses are endemic, eliciting a range of effects from a heavy burden of disease (e.g. malaria) to being essentially commensal in immuno-competent hosts (e.g. most types of herpesvirus and papilloma virus). Other infections depend on an animal reservoir for their maintenance; their infection of humans may be pathogenic, but it has little part in the evolving ecology of the microbe or parasite. An estimated 61\% of the 1,415 species of infectious organisms known to be pathogenic in humans are transmitted by animals\(^11\), for which the human represents a dead-end host. Occasionally, however, a zoonotic infection adapts to human-to-human transmission and diversifies away from its animal origin. Epidemic diseases are generally caused by infections that are directly transmissible between humans. HIV is a recent example of a long line of human infections initiated by a switch of host species, stretching back to the origins of measles and smallpox.

Free-living microbes may also find a human niche that suits their lifestyle, such as the lung for Legionella pneumophila and the gut for Vibrio cholerae. Legionnaires’ disease, first recognized in Philadelphia in 1976, is the environmental equivalent of a zoonosis. It is seldom passed directly from person to person but it was human ingenuity in designing warm, aerated, humid ‘artificial lungs’ called air-conditioning systems that allowed the microbe to proliferate and become an opportunistic colonizer of the human lung. Cholera, which was unknown beyond the Ganges delta before it spread widely in Asia and the Middle East during the period 1815–1825, at around that time

**Figure 1** Proportions of total deaths from major cause-of-death categories, 1909 and 1999, in Chile\(^2\). This country illustrates the full transition from developing to developed status during the twentieth century.
The values are approximate global death rates for the year 2003, taken from the World Health Organization (WHO) and other sources. HBV and HCV, hepatitis B and C viruses; RSV, respiratory syncytial virus; HPV, human papilloma viruses; vCJD, variant Creutzfeldt-Jakob disease.

Human society has undergone a series of major transitions that has affected our pattern of infectious disease acquisition and dissemination. These transitions illustrate the interrelationship between environmental, social and behavioral influences on the emergence and subsequent spread of infectious disease. Some infections were acquired when our australopithecine ancestors left their arboreal habitat to live in the savannah. This ecological change included exposure to new species of mosquito and tick as vectors for infection. After the emergence of Homo sapiens, the eventual migration of neolithic hunter-gatherers out of Africa 50,000 to 100,000 years ago exposed them to new infections in distant regions.

The first major transition of prehistoric/early historic times gave rise to the epidemic, or ‘crowd’, infections. This change must have started in the millennia following the advent of agriculture—from around 10,000 years ago—as agriculturally based society developed larger, denser populations. The domestication of livestock and the rich dividends available in human settlements to other animals (e.g., rodents, dogs and various insects) provided further opportunities for pathogens to move between species. Sometimes such a pathogen (or a mutant strain thereof) would have been well suited to humans as a new host species, and, if human numbers were adequate, could therefore persist indefinitely as a human infection. Thus, measles emerged about 7,000 years ago, probably from rundpest of cattle, and diverged to become an exclusively human infection when population size and density became sufficient to maintain the virus without an animal reservoir. Similarly, smallpox became epidemic about 4,000 years ago, possibly evolving from camelpox, its closest phylogenetic relative.

The next two transitions were primarily to do with great extensions in the spread of infectious diseases, entering distant populations as ‘new infections.’ Thus, the second historical transition occurred in Classical times as large Eurasian civilizations came into commercial and military contact. They inadvertently exchanged their pools of infections, and vectors such as rats and fleas, across the Mediterranean basin, the Middle East, India and China. The plague of Athens in 430 BCE during the Peloponnesian war vividly described by Thucydides may represent the first report of typhus. This Rickettsia infection is transmitted from rats to humans and thence among louse-ridden humans. Typhus frequently accompanies human conflict and deprivation, as seen in a recent outbreak among Rwandan refugees in Burundi. The Justinian plague of 542 CE devastated the eastern Mediterranean region and probably extended as far as China like the Black Death 800 years later (and both are attributable to Yersinia pestis).

The third historical transition accompanied the era of worldwide exploration and colonization by Europeans from circa 1500 CE onward. A contemporary account by one of Hernan Cortes’ fellow conquistadors, Bernal Diaz, recalls that they might well have failed to overthrow the mighty Aztec empire had they not been aided by a raging epidemic. This was possibly a combination of smallpox and measles, both wholly unknown to the New World population. Curiously, the Columbian exchange was unidirectional regarding infectious diseases; the one contentious possible exception being syphilis. The New World is believed to have had substantially fewer human zoonotic infections, and vector-borne infections like Chagas’ disease did not travel in the absence of an appropriate vector.

Two centuries later, Captain Cook unwittingly repeated the decimation of indigenous peoples through syphilis, measles and tuberculosis in many of the Pacific islands, whereas Lord Jeffrey Amherst deliberately attempted to spread smallpox among ‘hostile’ Native Americans, one of the better documented cases of germ warfare. The transmission dynamics of infections in naive populations is markedly different from those in which the majority of adults are immune. Onboard The Beagle, Charles Darwin observed with his customary acuity, "Wherever
the European has trod, death seems to pursue the aboriginal …Most of
the diseases have been introduced by ships and what renders this fact
remarkable is that there might be no appearance of the disease among
the crew which conveyed this destructive importation.”

Today we are living through the fourth historical transition of glob-
alization. Urbanization, dense and usually impoverished peri-urban
settlements, social upheaval, air travel, long-distance trade, technol-
gical developments, land clearance and climate change all influence the
risks of infectious disease emergence and spread. Although some of
the apparent increase in infectious disease may be attributable to bet-
ter diagnostic methods and surveillance, there seems little doubt that
more incidents are occurring, and have the potential to spread more
widely than 50 years ago, as outbreaks and spread of infections like
Nipah virus and SARS would not have passed unnoticed.

Environment and emerging infectious diseases
As humans encroach further into previously uncultivated environ-
ments, new contacts between wild fauna and humans and their
livestock increases the risk of cross-species infection. This process
will only diminish as wild species become rarer and eventually
endangered, like the great apes today. An example of such contact fol-
lowed the establishment of piggeries close to the tropical forest in
northern Malaysia, where, in 1998, the Nipah virus first crossed over
from fruit bats (flying foxes, Pteropus spp.) to pigs and thence to pig
farmers. Destruction of natural forest has also encouraged fruit bats
to relocate nearer human habitation, like the large colony in the botanic
gardens in the heart of Sydney. Indeed, in 1997, Hendra, a related
paramyxovirus of Australian fruit bats, fatally infected a veterinarian
examining a sick horse.

Rodents continue to be sources of re-emerging infections, as wit-
nessed in the 1990s with hantaviruses in the United States. Rodent-
borne hantavirus is prevalent in agricultural systems in South America
and East Asia, in arid grasslands in North America and elsewhere. In
mid-1993, an unexpected outbreak of acute, sometimes fatal, respira-
tory disease occurred in humans in the southwestern United States. This
‘hantavirus pulmonary syndrome’ was caused by a previously
unrecognized virus, maintained primarily within the native deer-
event, with unseasonal heavy summer rains and a proliferation of piñon
nuts, hugely amplified local rodent populations which led to the
1993 outbreak. In South America, there have been several out-
breaks of hantavirus and arenavirus infections linked to forest clear-
ance and the growth of rodent populations in the new grasslands.

Habitat destruction is not the only cause of increased human infec-
tion, however. Dengue virus is extending its range and prevalence
because its mosquito vector breeds rapidly in the urban environ-
ment. In the United States, nature conservation and increased wood-
land in the eastern states has led to the emergence of Lyme disease.
This disease is caused by a tick-borne spirochete and the presence of
tick-infested deer near suburban homes leads to ticks residing on
bushes adjacent to baseball diamonds and gardens.

Food-borne infections
Intensification of production of meat and meat products has led to
new infections. Most notorious is vCJD in the UK arising from
consumption of contaminated food products of cattle affected by
BSE. BSE, or ‘mad cow disease,’ emerged in British cattle in 1986
because of industrialized cannibalism, whereby rendered neural tis-
sue and bone meal from slaughtered cattle were recycled into cattle
feed, as well as into pies and hamburgers for human consumption.
Originally, infectious prions from scrapie in sheep were the suspected
source, but it now seems more likely that it arose from a bovine with
sporadic prion disease. The extent of the human epidemic remains
unclear. Although natural transmission is unsustainable (R₀ < 1 in
both cattle and humans), there are concerns that vCJD might be trans-
missible through blood transfusions. Without effective diagnostic
tests for presymptomatic vCJD infection, this situation is extremely
unfortunate.

Other recently emergent food-borne infections include E. coli
O157:H7, which is harmless to cattle but toxic to humans, and
Salmonella enteriditis in chickens. Better hygiene in abattoirs, butchers
and domestic kitchens can greatly reduce the incidence of infection. In
theory, closed and intensive farming of a single species should reduce
the risk of cross-species infection (Fig. 3). But it also allows large-scale
epidemics to emerge, as seen recently for avian influenza strains in
southeast Asia and the Netherlands.

Ancient dietary taboos, such as those of Hindus, Muslims and Jews
regarding pork as unclean, doubtless had their roots in protection
from infectious disease. Today, an increasing demand for consumption
of exotic and wild animals raises new risks of infectious diseases such
as SARS (Box 2).

Transmissibility and globalization
Changing patterns of human behavior and ecology affect two distinct
steps in the emergence of new infectious disease. The first is an

Figure 3 The changing pattern of farming in South East Asia. Top,
traditional mixed homestead; bottom, intensive single-species industry. (Top
photo courtesy of R.A. Weiss)
increased opportunity for animal-to-human infection to occur owing to greater exposure, which may be necessary but not sufficient to lead to the emergence of a new human infection. The second step is the opportunity for onward transmission once a person has become infected. For each novel epidemic, such as the 1918–1919 influenza pandemic or AIDS, there are probably thousands of failed transfers.

Some infections simply do not take in the new host. Innate host-specific restrictions on viral replication have recently become evident for primate lentiviruses31, which may explain why certain species that harbor simian immunodeficiency virus, but not others more commonly in contact with humans, gave rise to HIV-1 and HIV-2. Even in the case of HIV-1, only one pedigree of three independent chimpanzee-to-human crossover events32 has given rise to the AIDS pandemic, whereas the other two smolder as poorly transmissible infections.

Fatal pathogenesis is not necessarily coupled with infectiousness32, which is evident for H5N1 avian influenza in humans8. But genetic reassortment between avian and human influenza viruses could easily give rise to a new, rapidly spreading strain8. A poorly infectious pathogen may not spread at all from the index case, as is usual with rabies, or may only infect close contacts and soon peter out, as seen with Lassa fever and Ebola virus. SARS nearly became self-sustaining but was brought under control. Some of the most insidious infections are those with long, silent incubation periods during which the person is infectious. These emerge surreptitiously so that when the new disease is eventually recognized, as AIDS was in 1981, the infection has already spread far beyond control.

Like the ships of centuries past, the speed of modern air travel works wonders for the dispersal of infectious diseases. SARS was eventually constrained by quarantine and strict adherence to infection control guidelines in hospitals, but not before it quickly traveled from Guangzhou to Hong Kong and on to Toronto. If Ebola broke out in a city with a busy international airport, it might also travel across continents in a similar manner. Brockmann33 has modeled how rapidly such infections can move once they reach a major airport hub; closing the hubs becomes an immediate imperative. We cannot be sure what the initial vector was for the arrival of West Nile virus into North America in 1999: a migratory bird blown off course, an infected human with a valid air ticket or a stowaway mosquito on a similar flight. Whatever the means of entry and early colonization of crows in New York, it has taken less than four years to reach the Pacific coast25. Thus, West Nile virus has found a new reservoir in American birds, just as yellow fever virus reached New World primates 350 years earlier.

Social and economic conditions, behavioral changes and geopolitical instability

Microbes frequently capitalize on situations of ecological, biological and social disturbance. Biologically weakened and vulnerable populations—especially if also socially disordered and living in circumstances of privation, unhygienic conditions and close contact—are susceptible to microbial colonization. The severity of the bubonic plague (Black Death) in mid-fourteenth-century Europe seems to have reflected the nutritional and impoverishment consequences of several preceding decades of unusually cold and wet weather with crop failures compounding the incipient destabilization of the hierarchical feudal system.

Many of the rapid and marked changes in human social ecology in recent decades have altered the probabilities of infectious disease emergence and transmission. These changes include increases in population size and density, urbanization, persistent poverty (especially in the expanding peri-urban slums), the increased number and movement of political, economic and environmental refugees, conflict and warfare. Political ignorance, denial and obduracy often compound the risk of infectious disease transmission—as has been tragically observed with HIV/AIDS in parts of Africa, where widespread poverty, a culture of female disempowerment and political instability further
exacerbate the problem\textsuperscript{34,35}. But we have little understanding of why the prevalence of HIV infection varies so greatly between cities in sub-Saharan Africa\textsuperscript{36}.

The urban environment has only recently become the dominant human habitat. Urbanism typically leads to a breakdown in traditional family and social structures, and entails greater personal mobility and extended and changeable social networks. These features, along with access to modern contraception, have facilitated a diversity of sexual contacts and, hence, the spread of sexually transmitted diseases\textsuperscript{37}. This risk is further amplified by the growth in sex tourism in today's internationally mobile world, which capitalizes on the desperation and ignorance of poverty, combined with exploitative behaviors, in developing countries. More generally, cities often function as highways for ‘microbial traffic’\textsuperscript{38}. Rapid urbanization boosts certain well-established infectious diseases, such as childhood pneumonia, diarrhea, tuberculosis and dengue, and facilitates dissemination of various ‘emerging’ diseases—as occurred for SARS in the high-rise housing of Hong Kong. Crowded and dilapidated public housing can potentiate infectious disease transmission through drug abuse and sexually transmitted infections\textsuperscript{39,39}.

**Nosocomial and iatrogenic infections**

Technological advances in medicine and public health can also inadvertently promote the emergence and spread of infectious disease. It has become commonplace to quip that you go to the hospital at the peril of acquiring an intractable nosocomial infection such as methicillin-resistant *Staphylococcus aureus*\textsuperscript{40}, and such infections killed around 40 times as many people as SARS did in 2003 (Box 1). Multidrug-resistant tuberculosis has also become a major problem, and, paradoxically, regions with health programs that reduced wild-type tuberculosis strains can develop into ‘hot zones’ for multidrug-resistant tuberculosis\textsuperscript{41}.

By far the most effective medical vector of infectious disease has been the syringe and needle. Drucker et al.\textsuperscript{42} have charted the massive increase in the use of injecting equipment over the past 100 years. Individuals with hemophilia treated with pooled clotting factors became almost universally infected with hepatitis B and C viruses before diagnostic screening tests were developed. Over 20% of such affected individuals also became infected with HIV\textsuperscript{43}, and more recently, transmission of West Nile virus by blood transfusion and by organ transplantation has been reported\textsuperscript{44,45}. The use of contaminated needles among intravenous drug users has had similar consequences. Infectious diseases have also been amplified by the use of nonsterile medical injections in developing countries\textsuperscript{42}.

Egypt has the highest prevalence of hepatitis C infection in the world because of the use and reuse of syringes and needles in an earlier public health campaign to reduce bilharzia by medication given by injection. The transmission of CJD through contaminated surgical instruments is another example of iatrogenic spread of infection\textsuperscript{46}.

Biological medicines produced from animal-cell substrates present an inherent potential hazard for introducing new infections. Great care must be taken to ensure that live attenuated vaccines grown in animal cells or eggs are devoid of pathogens; for example, several early batches of live and inactivated polio vaccine unwittingly contained live SV40 virus, a polyoma virus of macaques. After SV40 was discovered in 1960, polio vaccine production shifted to virus propagation in primary kidney cells of African green monkeys. These cultures were free of SV40 but possibly contained SIVagm, a relative of HIV that fortunately does not infect humans\textsuperscript{47}. The irony of the SV40 story is that the United States Food and Drug Administration prohibited the use of well known, permanent cell lines demonstrably free of adventitious infectious agents, for fear that such immortalized cells might exert oncogenic properties on the vaccine. There is no epidemiological evidence of increased tumor incidence in those populations who are known to have received SV40-contaminated polio vaccine. But there have been a number of recent claims of an association of SV40 DNA sequences in a variety of human malignancies\textsuperscript{48}, although these findings remain controversial\textsuperscript{49}.

The ultimate medical means of introducing animal viruses into humans is xenotransplantation. The implantation of animal cells or tissues into immunosuppressed individuals seems to be a perfectly designed way to encourage cross-species infection. It is astonishing that trials were started without much thought about the consequences for potentially emerging pathogens, for example, porcine retroviruses\textsuperscript{50}. The generation of genetically modified knockout or transgenic animals to prevent hyperacute rejection of donor tissues may exacerbate the infection hazard\textsuperscript{49,50}. Happily, there is no evidence so far of retrovirus infection in individuals who were exposed to living pig cells\textsuperscript{50}, and clinical xenotransplantation is now stringently regulated; so it seems all the more extraordinary that cellular therapies with fetal lamb cells and extracts continue to be practiced with impunity in alternative medicine clinics in Europe and the Far East.

**Conclusions and prospects**

Novel infectious diseases can emerge in any part of the world at any time. HIV and Ebola came out of Africa, avian influenza and SARS from China, Nipah virus from Malaysia, BSE/ vCJD from the UK and hantavirus pulmonary syndrome from the Americas. It is difficult to predict what new disease will come next or where it will appear, but changing ecological conditions and novel human-animal contacts will be useful clues as to which horizons require scanning with most scrutiny. We must expect the unexpected.

As a codicil, another factor that needs to be taken into account is the potential impact of the HIV pandemic on the emergence of other infectious diseases\textsuperscript{51}. We already know that persons with AIDS act as ‘superspreaders’ of tuberculosis, and we can only speculate what course the SARS outbreak might have taken had someone incubating the disease flown to Durban rather than Toronto\textsuperscript{52}. People with AIDS may persistently harbor infections that would otherwise be transient, and this could hamper the eradication of measles and polio. Multivalent *Pneumococcus* vaccines are ineffective in HIV-infected people with CD4\textsuperscript{+} lymphocyte levels below 200/\mu l, whereas live ‘attenuated’ vaccines such as vaccinia can cause virulent disease in the immunocompromised host. Immunodeficient persons living at high density could also be the seed-bed for microorganisms that are initially ill adapted to human infection to evolve into transmissible human pathogens. Thus, an infection from a zoonotic or environmental source—for example, the *Mycoplasma avium intracellulare* complex—could conceivably emerge as the tuberculosis of the twenty-first century, although direct transmission between individuals with AIDS of such opportunistic infections have not been documented so far.

We shall give Girolamo Frascatoro the last word on emerging and re-emerging infectious diseases by quoting from his treatise *De Contagione*, published almost 450 years ago, “There will come yet other new and unusual ailments in the course of time. And this disease [syphilis] will pass way, but it later will be born again and be seen by our descendents.”

**Acknowledgments**

We are grateful to M. E. Chamberland, H. W. Jaffe and S. Leff for critically reading the manuscript.