Evolutionary biology has occupied a distant position within the health sciences. Evolution is widely recognized as providing the ultimate foundation for biology and hence for the health sciences and is widely accepted as being of practical importance in the explanation of antibiotic resistance and phylogenetic analyses of pathogen strains. Yet, evolutionary processes are often dismissed as being of little practical importance for prospective study even though pathogens, with their short generation times, can evolve predictably within periods as short as a few years, months, or weeks. Even in the case of antibiotic resistance, evolution is looked upon more as a reason for a problem than as a process that can be controlled as part of the solution.

An accurate evolutionary approach should unite rather than replace the insights from mechanistic nonevolutionary approaches. With this goal in mind, we offer an evolutionary approach to disease causation and use this perspective to address both the possible causative role of Chlamydia pneumoniae in chronic diseases and practical applications that may result from improved understanding of this role. We begin by emphasizing the general explanations for disease causation and how evolutionary considerations may help distinguish their relative roles.

Disease Causation and Evolutionary Fitness

Diseases can be attributable to genetic, infectious, and noninfectious environmental causes. Although these categories could in principle be mutually exclusive, all infectious diseases that have been adequately studied are also influenced by host genetics and noninfectious environmental factors. The corollary of this fact is that the identification of genetic or noninfectious environmental influences on disease cannot be used as evidence against infectious causation. Such influences are expected among infectious diseases.

The presence of common influences, however, does not mean that each factor has a similar role in the hierarchy of disease causation. Some risk factors will be primary causes that initiate the disease process; others will be secondary causes that accelerate the disease process or exacerbate the damage. Still other risk factors may not play a causal role but may simply be correlated with causal factors. As the attention of medicine shifts to chronic diseases with complex causal mechanisms, these distinctions become increasingly important. If the distinctions are ignored, causal hypotheses and hence causal mechanisms may be overlooked as researchers engage in the slow process of determining the extent to which different risk factors are associated with disease. Although causal thinking has been central to scientific progress, the concern for objectivity and perhaps the fear of criticism seem to have led researchers to back away from hierarchical causal arguments, replacing them with the safer allocations of risk to various variables. We believe that this tendency is detrimental to scientific understanding because failure to formulate hypotheses of hierarchical causation hinders the tests of such hypotheses and hence the discovery of primary causes.

In this context, we define primary causes as those that are necessary for the disease in question. Elimination of the primary cause (or causes) will eliminate the disease. If the primary cause is an infectious agent, the associated disease is appropriately referred to as an infectious disease, even though genetic and noninfectious environmental factors may make important contributions. If, for example, Mycobacterium tuberculosis is eliminated, tuberculosis will not occur. M. tuberculosis is therefore considered the primary cause of tuberculosis, and tuberculosis is considered an infectious disease, even though human genetic variation and noninfectious environmental factors may influence the expression of M. tuberculosis infections.

The concept of evolutionary fitness provides useful clues about disease causation. This concept has been central to evolutionary theory from Darwin’s time to the present. Before the
establishment of genetics, evolutionary biologists used the concept to refer to the fit between the organism and its environment; improved fit was favored because it resulted in increased propagation into succeeding generations through increases in survival and reproduction. With the establishment of genetics, fitness has become equated with genetic contribution into the succeeding generations. When applied to genes, the fitness of an allele refers to the change in its representation over time relative to alternative alleles.

The fitness concept can be applied to the problem of disease causation to distinguish evolutionarily feasible hypotheses of causation from marginally feasible or untenable ones [1]. Specifically, if the negative effects of a disease are so great that genetic instructions for the disease could not be maintained through time by mutation, then strict genetic causation can be rejected as untenable. Monozygotic twin concordance serves as a check; low-to-moderate concordance indicates that some environmental cause, either infectious or noninfectious, is playing a major role. If a disease distribution in time and space makes a common noninfectious cause an unlikely explanation of the remaining variation in incidence, then attention is directed to infectious causation. These points are summarized diagrammatically in figure 1.

Although the three categories of disease causation are not mutually exclusive, infectious causation has often been accepted belatedly throughout the history of medicine. The delay in recognizing infectious causation is attributable to characteristics of the organism, the infection, and transmission that make infectious causation cryptic relative to the investigative tools of the day. This effect continues to the present day and is manifested in the current controversy over infectious causation of atherosclerosis [1].

**Evolutionary Implications of Infectious Etiology of Atherosclerosis**

Application of the evolutionary approach to atherosclerosis implicates infectious causation. The effects of atherosclerosis on myocardial infarction and stroke occur largely after direct reproduction; they therefore probably have had their major effect on fitness through reductions in resources available to children and other relatives such as grandchildren. We estimate the fitness reduction due to myocardial infarction and stroke to be about 1% [1]. Although this figure is a rough estimate, it is about two orders of magnitude above the percentage that could be maintained by mutation if these diseases were primarily the result of a genetic predisposition.

The maximal genetic basis for a disease can be gauged for a particular environmental context by the monozygotic twins concordance, which quantifies the probability that a person will have a disease when the disease occurs in that person’s monozygotic twin. The monozygotic twin concordance for coronary heart disease, for example, is about 50%, roughly the same as for tuberculosis [2]. This value supports the fitness-based argument for an important causative role by something other than host genetics.

![Diagram](image)

**Figure 1.** Three categories of disease causation. Location within triangle corresponds to relative importance of genetic, infectious, and noninfectious environmental causation of disease—increased distance from an apex indicates reduced importance of the designated cause. Arrows A and B designate collections of factors that decrease feasibility of genetic and environmental explanations and by default implicate infectious causation. Evidence of reduced importance of genetic causation includes long history of high fitness costs of disease averaged over the host population, low monozygotic twin (mzt) concordance, and changes in disease frequency too rapid to be explained by genetic change. Evidence of reduced importance of noninfectious environmental influences include long history of disease, wave-like spread of disease through host population, and cycling of disease over time; each process commonly occurs among infectious diseases but would be unusual for noninfectious environmental causes.
The presence of atherosclerosis in Egyptian mummies [3] indicates that the preservable manifestations of atherosclerosis have a long history. Long-present, noninfectious environmental factors (e.g., high iron and lipid intake) are therefore unlikely primary causes because over such long periods of time natural selection probably would have reduced the vulnerability to such constituents.

In contrast with noninfectious environmental causes, infectious causes set up evolutionary arms races between host and pathogen and can therefore be associated with high levels of damage to the host indefinitely over evolutionary time [4]. This line of reasoning narrows the spectrum of likely candidates for primary causation of atherosclerosis to new noninfectious environmental influences (e.g., smoking) and infectious causation. This conclusion does not resolve the central issue of this symposium, namely whether C. pneumoniae in particular causes atherosclerosis. But by implicating infectious causation as one of the few feasible primary causes of atherosclerosis, evolutionary considerations lend support to the argument that research should focus on the various possible infectious etiologies [5].

Genetic Predispositions and Infection

Arguments about loss in evolutionary fitness can be applied to known alleles and to disease entities. The ε4 allele of the apolipoprotein E (ApoE) locus, for example, is associated with increased risk of atherosclerosis, stroke, Alzheimer’s disease, and severe cases of multiple sclerosis [6–10]. The fitness reduction associated with this set of diseases is so high (i.e., >1%) that the ε4 allele could not be maintained without counterbalancing benefits.

One could argue that this fitness loss is offset by some other pleiotropic benefit of the ε4 allele, but this argument narrows substantially the scope for strict genetic causation. Compensatory pleiotropic benefit from disease-causing alleles may involve protection against infectious diseases; if so, the perpetuation of such diseases again depends on an arms race with infectious agents. That is, infectious causation is implicated in an evolutionary sense. One well known example is the protection against malaria afforded by the allele for sickle cell anemia. Pleiotropic benefits other than protection against infectious disease are possible but have been not been documented for common damaging diseases, such as atherosclerosis. Genetic vulnerability to infection, however, appears to be a pervasive cause of allelic associations with disease [1, 11].

Research on arthritis patients suggests that the association between the ε4 allele and chronic disease results from this mechanism: The ε4 allele is associated with increased risk of C. pneumoniae infection [12]. This finding draws together a great deal of seemingly unrelated or discordant results and suggests a variety of directions for further study.

First, from an evolutionary perspective it offers explanations for the maintenance of the ε4 allele over time. Because pathogens are continually evolving with their hosts, the genetic basis for vulnerability to a pathogen is likely to change over time. Alleles conferring vulnerability (e.g., the ε4 allele) may be presently disfavored and declining relative to alternative alleles. But if the vulnerability conferred by the ε4 allele allows C. pneumoniae infections to become more productive, a declining frequency of ε4 alleles over time would increase the selection pressure on C. pneumoniae to evolve abilities to exploit humans with alternative alleles, thus generating a collection of alleles, each of which may confer vulnerability at a particular time depending on the characteristics of C. pneumoniae at that time. This argument is a special case of “dynamic polymorphism,” which can be maintained over time by “nonprogressive seething” [13, 14]. Alternatively, the ε4 allele could have been maintained if C. pneumoniae was restricted in space for much of human history, and the selective pressure against the ε4 allele were thus reduced. Or, C. pneumoniae could have entered humans relatively recently, too recently for it to have eliminated the ε4 allele.

With regard to the causes of atherosclerosis, the important point is not whether any one of these hypotheses is more likely than any other. Rather it is that the ε4 allele could be maintained indefinitely in a population despite the fitness costs it incurs so long as the costs are incurred through increased vulnerability to an infectious agent such as C. pneumoniae.

This line of logic helps explain seemingly discordant findings about suites of related diseases. Different forms of Alzheimer’s disease, for example, are associated with different risk factors. Early onset Alzheimer’s disease is a genetic disease in the strict sense; it is sufficiently rare to be explained by mutation rate and has a high monozygotic twin concordance. The ε4 allele does not, however, appear to be a risk factor for early onset Alzheimer’s disease [15]. The more common sporadic Alzheimer’s disease has a monozygotic twin concordance that is well within the range of typical infectious diseases and has been associated with C. pneumoniae infection [12]. The ε4 allele rather than atherosclerosis per se is implicated in this form of Alzheimer’s disease [16], raising the possibility of a direct effect of C. pneumoniae on Alzheimer’s disease rather than an indirect effect via atherosclerotic damage. The available evidence therefore accords with the classification of the two categories of Alzheimer’s disease as two fundamentally different diseases that show some phenotypic similarities, one being a genetic disease and the other caused by infection. Like other infectious diseases, the latter is associated with genetically based variation in susceptibility to infectious damage and has C. pneumoniae as the leading candidate for the agent of this damage.

The possible causative role of C. pneumoniae has been studied much less intensively in Alzheimer’s disease than in atherosclerosis. Yet, as with atherosclerosis, the current state of evidence on infectious causation is mixed [17]. The mixed evidence must result from differences in procedures (e.g., in study design,
subjects, protocol, and execution). Future work should clarify whether the negative or the positive results provide the more accurate indication of the processes involved, and hence whether the conceptual cohesiveness provided by the C. pneumoniae model of causation is real or illusory.

The emphasis on infectious causation and the genetic variability in susceptibility may also help clarify some confusion over infectious causation of other members of the suite of diseases that may be influenced by C. pneumoniae and the e4 allele. Epidemiologic patterns, fitness costs, and the low monozygotic twin concordance associated with multiple sclerosis, for example, implicate infectious causation [18]. A recent report of an association of multiple sclerosis with C. pneumoniae [19] has generated skepticism and a call for independent confirmation [18]. This process of verification may be more efficient by drawing upon the full spectrum of information about e4 and C. pneumoniae. If C. pneumoniae is causally involved and the e4 allele increases vulnerability to C. pneumoniae, then one would expect an association between the e4 allele and multiple sclerosis.

Without reference to C. pneumoniae, a recent study of ApoE alleles among multiple sclerosis patients found that the e4 allele is not associated with unselected multiple sclerosis patients, with relapsing cases, or with primary progressive patients; however, the allele was associated with particularly aggressive cases of multiple sclerosis [8]. This set of results is therefore consistent with C. pneumoniae exaceriating a disease process that is set in motion by other infectious agents. This interpretation draws attention to the need to replicate each of these studies and also suggests where to look for an association between C. pneumoniae and multiple sclerosis—in persons with severe multiple sclerosis who also have the e4 allele.

These arguments offer a broad causal perspective on e4-associated chronic diseases (figure 2). Rather than viewing e4 as a deleterious allele that is responsible for damaging cardiovascular and neuronal tissue, this new perspective casts the e4 allele as an Achilles heel that makes a person vulnerable to C. pneumoniae infection. This argument thus casts C. pneumoniae infection as a primary cause of e4-associated diseases.

The evidence for a connection between the e4 allele and C. pneumoniae also offers an explanation for apparent contradictory associations between the e4 and cardiovascular disease in different geographic areas [20, 21]. The different associations may result from differences in the prevalence of infectious agents. South African blacks, for example, have a higher frequency of e4 but a lower frequency of atherosclerosis than do residents of the United Kingdom (UK); this apparent contradiction might result if C. pneumoniae is less pervasive in South Africa than in the UK (e.g., as a result of climatic differences that influence C. pneumoniae transmission). Comparisons among people of different genetic makeup in the same living environment are needed to sort out effects of ApoE alleles, other genetic effects, and noninfectious environmental influences, each of which might have effects on atherosclerosis through effects on C. pneumoniae infection.

**Interplay of Environmental Factors and Infection**

As indicated by the evidence from ApoE alleles, noninfectious risk factors may not only be consistent with infectious etiologies but difficult to explain without invoking infection (see also [22, 23]). Smoking, for example, is a risk factor for atherosclerosis [24, 25], may aggravate pulmonary infection, and is associated with increases in immunologic indicators of C. pneumoniae infection [22, 26]. The exacerbation of C. pneumoniae infections by smoking may contribute to atherosclerosis by increasing the potential for systemic establishment of C. pneumoniae infection.

Exposure to smoke from smokers (termed passive smoking) has also been implicated as a risk factor for atherosclerosis [27] and atherosclerosis-associated diseases such as stroke [28]. This exposure is associated with an increased risk of about one-third of the increased risk associated with smoking [27]. This relative effect of passive smoking seems out of proportion to the small amount of smoke inhaled relative to that inhaled by smokers, a point that has been made in criticism of the study’s conclusions [29]. If the reported effects of passive smoke are correct, the dose-effect curve must rise very sharply at very low doses of smoke. From an evolutionary perspective, this extreme sensitivity to low doses of smoke seems unlikely for humans, who have spent most of their evolutionary history in smoky environments. Effects of exposure to second-hand smoke on antioxidants and lipid metabolism are interpreted as a mechanism by which second-hand smoke may cause heart disease [30].
These effects, however, are consistent with infectious causation. They could be part of the defense mechanism against smoke or nonadaptive side effects that cause little if any chronic damage unless the system is compromised by the lipophilic *C. pneumoniae*.

Exposure to the more florid or more frequent infections of smokers is an alternative hypothesis for the effects of passive smoke. This hypothesis seems especially feasible because it needs to assume only that the infection-proneness of smokers (rather than the small amount of smoke itself) would increase the risk of *C. pneumoniae* transmission from them by one-third. Associations between exposure to second-hand smoke and increased frequencies of infection have been documented for another bacterial pathogen of the respiratory tract, *Haemophilus influenzae* [31–33]. Similar studies of *C. pneumoniae* could help clarify whether exposure to *C. pneumoniae* from smokers is a feasible causal explanation for the risk attributed to passive smoking. The studies of *H. influenzae* did not distinguish between two explanations for the increased infection rate among nonsmokers: The increase could have resulted from compromised antibacterial defense mechanisms due to passive smoking or from the increased exposure to *H. influenzae* from the more florid infections of smokers. These two possibilities would need to be considered for a thorough assessment of effects of passive smoke on *C. pneumoniae* infection and cardiovascular disease.

The consideration of ApoE alleles and smoking provides two illustrations of a comprehensive approach to the causation of atherosclerosis. Instead of accumulating risk factors, this approach seeks a unified theory of causation. Infection as a primary cause of atherosclerosis provides a conceptually cohesive framework for understanding the role of noninfectious environmental risk factors, such as smoking and diet [5, 34]. This integration of risk factors was recently extended to lipid metabolism with evidence that *C. pneumoniae* induces cellular oxidation of low-density lipoproteins and lipid accumulation [35] (Byrne, elsewhere this issue). The association between high iron levels and atherosclerosis [36] may similarly be explained by the protective effects of low iron levels against bacterial infection [37]. Associations between atherosclerosis and indicators of immunologic activity (e.g., C-reactive protein and inflammation) may occur because microbes are activating the immune system. Hyperlipidemia and obesity may act in concert with the lipophilic *C. pneumoniae* to cause excessive fat deposition in atherosclerotic plaques. Other risk factors, such as hypertension, may be consequences of cardiovascular damage rather than a primary cause.

The intriguing possibility raised by the e4 associations is that even this broad conceptual integration of risk factors for atherosclerosis may not be broad enough. The complete picture may encompass similar integrations of risk factors for other chronic conditions such as Alzheimer’s disease.

### Evolution of Antibiotic Resistance

The current antibiotic resistance problem is largely a result of evolutionary processes. We can therefore expect that improved integration of evolutionary insights might suggest improvements in the control of antibiotic resistance. From the middle of the twentieth century to the present, the health sciences have responded to the problem of antibiotic resistance rather than considering how policy could be planned prospectively and strategically to prevent problems of antibiotic resistance before they arise.

Given the existing uncertainties about the role of *C. pneumoniae* in atherosclerosis, it may seem premature to consider the control of cardiovascular damage by controlling the resistance of *C. pneumoniae* to antibiotics. If, however, this topic is postponed until the role of *C. pneumoniae* in cardiovascular disease is resolved, an opportunity for preserving antibiotic efficacy may be lost. Control efforts may be able to capitalize on a lucky break, namely that the most damaging infections may be dead-end infections; if so, if effective antibiotics are developed for treatment of these infections and if appropriate guidelines of antibiotic usage are enacted, these damaging infections may be controlled indefinitely by antibiotics without generation of antibiotic resistance. For this possibility to be realized, research needs to resolve two uncertainties before *C. pneumoniae* is shown to be a cause of atherosclerosis.

The first uncertainty concerns the degree to which the *C. pneumoniae* that cause systemic infections are transmissible to other people (figure 3). Age-related seropositivity suggests that systemic *C. pneumoniae* may be relatively or wholly nontransmissible. Seropositivity is very low among children who have not yet attended school, rises steeply during the next three decades of life, and then levels off in older age groups, who remain predominantly seropositive [38–42]. About one-third of the adults who are in the typical age range of parents of such children (e.g., 25–35 years old) have some atherosclerotic damage [43]. If *C. pneumoniae* causes atherosclerosis and if these systemic infections were readily transmissible, one would expect to see a much higher frequency of infection among children who are not yet school age as a result of transmission from their atherosclerotic parents. The fact that young children are sometimes infected at a higher frequency [44] argues against an age-dependent barrier to transmission. Still, one could hypothesize that these trends result from low rather than no transmissibility of systemic infections or by some characteristic of children, such as partial resistance to infection or incomplete immunologic responses. Prospective studies are therefore necessary to assess directly the degree of transmission from systemic infections.

Enactment of wise guidelines for antibiotic treatment of pulmonary and systemic infections will depend on the resolution of this matter. If systemic infections are not transmissible, then the efficacy of antibiotics that are effective against systemic
infections may be preserved indefinitely by keeping them from being used for treatment of the transmissible pulmonary infections. If such antibiotics are used only for systemic infections, we can expect some resistance to evolve within some treated persons. But if systemic infections are not transmissible, these antibiotic-resistant variants would not be transmitted, and the antibiotic resistance would not have a chance to increase progressively from infection to infection. Rather, when the infection is cleared or the person dies, any antibiotic resistance that had evolved would be lost, and the evolution of antibiotic resistance would need to begin anew. This strategy for controlling antibiotic resistance can therefore be termed the Sisyphean control in reference to the mythic Greek, Sisyphus, who was condemned by Zeus to roll a rock uphill and forever to restart the task before reaching the hilltop.

If, however, such antibiotics are used indiscriminately for treatment of both pulmonary and systemic infections, we can expect the now familiar stepping stone development of antibiotic resistance: The resistance that evolves during the treatment of pulmonary infections can increase indefinitely, thus negating the efficacy of treatment of both pulmonary and systemic infections. The result would be that the most important antibiotics—those that provide effective treatment of the damaging systemic infections—would lose their effectiveness through the evolution of antibiotic resistance.

If systemic infections were just as transmissible as pulmonary infections, then the separation of antibiotics by category of infection would not provide greater long-term control. If systemic infections are transmissible, but less so than pulmonary infections, then the best strategy would involve more guesswork, but separating antibiotics according to systemic and pulmonary infection would likely provide at least some improvement in long-term control of antibiotic resistance in systemic infections.

Thus, determination of the appropriate guidelines for antibiotic treatment will depend on knowing the degree to which systemic infections are transmissible. This knowledge about transmissibility is most useful if it is acquired before the role of C. pneumoniae in atherosclerosis is resolved and before the discovery of antibiotics that are efficacious against systemic infections. As soon as a causative role for C. pneumoniae is demonstrated, we can expect a major increase in antibiotic treatment of pulmonary infections in order to reduce chances of developing systemic disease. Any antibiotics that are effective against systemic infection would need to be recognized as such by this time so that their efficacy can be preserved by restricting their use to treatment of systemic infection. Physicians cannot be expected to withdraw a particular antibiotic for pulmonary infection based on the speculation that the antibiotic might be effective in systemic treatment; if it is known that the antibiotic is effective in systemic infections and that systemic infections are not transmissible, the argument for withholding of the antibiotic would be compelling, so long as there are other antibiotics available for pulmonary treatment.

This last point identifies the other critical need: the development of a panel of effective antibiotics that includes at least one antibiotic that is just as effective against pulmonary infection as the antibiotic that is to be reserved for systemic infection. The antibiotics for systemic and pulmonary treatment should have different mechanisms of action so that resistance against the antibiotic for pulmonary infection would not generate cross-resistance against the antibiotic for systemic infection.

The difficulty in eradicating C. pneumoniae through antibiotic treatment of pulmonary infections [45] further emphasizes the importance of resolving this issue. When pathogens are reduced rather than eliminated by antibiotic treatment, the risk of evolving antibiotic resistance through a stepping stone process is increased. Antibiotic resistance might therefore evolve rapidly in response to extensive treatment of pulmonary infections.

Conclusions

The various risk factors for atherosclerosis are often seen simply as alternatives to each other and to infectious causation, with each alternative perhaps making a particular and relatively
infectious contribution to the disease. An evolutionary approach to cardiovascular disease implicate infectious causation and provides the basis for integrating the evidence for infectious and noninfectious risk factors into a unified comprehensive perspective as is suggested for the influence of the risk factors such as ApoE alleles, iron levels, and smoking on cardiovascular and neurologic diseases.

Evolutionary considerations of antibiotic resistance and transmissibility raise two goals that need to be met soon. The first is to determine whether C. pneumoniae is transmissible from systemic infections. This needs to be understood before C. pneumoniae’s role in atherosclerosis is resolved and before efficacy of treatment of systemic infections is demonstrated. The second goal is to develop a spectrum of antibiotics that is effective against pulmonary infections so that antibiotics found to be effective in treating systemic infection can be reserved for systemic infections.

Acknowledgments

The ideas presented benefited from communications with J. T. Grays-ton, W. D. Hamilton, Alan Hudson, F. J. Nieto, and an anonymous reviewer.

References


