Coronary Heart Disease (CHD)—One or Several Diseases?

changes in the prevalence and features of CHD

Maria Inês Azambuja* and Richard Levins†

ABSTRACT  In retrospect, mortality from coronary heart disease (CHD) in the 20th century followed an epidemic pattern: mortality rates increased dramatically from 1920 until about 1960, remained roughly constant for almost a decade, and have been decreasing since the late 1960s. CHD has traditionally been conceived of as a single disease with multifactorial causality. We suggest instead that CHD cases may comprise at least two distinct populations: those associated with hypercholesterolemia, and those associated with insulin resistance. The epidemic of CHD was due primarily to changes in the incidence of the hypercholesterolemia subgroup. We propose that young adults who survived the 1918 influenza pandemic were rendered vulnerable to lipid-associated CHD and coronary thrombosis upon reinfection with influenza later in life. This vulnerability may be due to autoimmune disruption of low-density lipoprotein-receptor interactions. Historical events may affect the health of populations by affecting the susceptibility of populations to chronic diseases such as CHD. The life experiences of individuals are known to influence their susceptibility to infectious diseases; we suggest that life experiences may also influence individual susceptibility to chronic diseases.

*Graduate Program of Epidemiology, School of Medicine, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil. To whom correspondence should be addressed.
†School of Public Health, Harvard University.
E-mail: miaazambuja@terra.com.br.

Perspectives in Biology and Medicine, volume 50, number 2 (spring 2007):228–42
© 2007 by The Johns Hopkins University Press
CORONARY HEART DISEASE (CHD) refers to “cardiac events or symptoms related to myocardial ischemia and/or injury due, in the vast majority of cases, to atherosclerosis. Such events include unstable angina, myocardial infarction (MI), and sudden death due to ischemic heart disease” (Pasternack et al. 2003, p. 1863). CHD mortality rose after World War I and attained public health significance after World War II (Morris 1951; Stallones 1980a). In the United States, the number of deaths attributed to the disease increased until the early 1960s, when it leveled off at rates of 35% of the overall mortality (Havlick and Feinlieb 1979). An unexpected decline in CHD death rates began in 1968 and accelerated after 1972, resulting in a fall of more than 60% in CHD deaths in the last 35 years (Levy 1981; NHLBI 2002). Similar trends have been documented in other countries (Thom 1989).

This article revisits evidence that may have been overlooked or misinterpreted during the heights of the CHD epidemic due to constraints posed by the idea of “degenerative” diseases, the methodologic framework of multicausality, and the hegemony of the lipid over alternative hypotheses to CHD causation (Armstrong, Conn, and Pinner 1999; Marmot 1976; Morris 1964; Pearce 1996; Stephens 1992; Susser and Susser 1996; Taubes 2001). A fresh look at some old evidence leads to new ways of thinking about CHD, its trends and its causes, and new ways of thinking about chronic disease occurrence in general.

The Degenerative Paradigm

CHD mortality increased at a time when economic and public health achievements seemed to have eliminated most of the major causes of infection-related deaths in developed countries. Explanations for the increase in CHD followed the common sense of the times: besides population aging, causes should be sought among lifestyles associated with urbanization and economic growth, assumed to favor chronic-degenerative processes (Armstrong et al. 1999).

The degenerative paradigm has had a strong influence in the way we conceive causality and patterns of disease in populations. The idea of degeneration implies that each individual experiences the same law-like history during his or her lifetime: “As a group, seventy-year-olds are grayer and more forgetful than thirty-five-year-olds because all individuals have been aging in body and mind” (Lewontin 2001, p. 8). This idea is in accordance with a Platonic understanding of nature, which conceives of individuals as adequately represented by their average presentation, and variation as accidental realizations of an idealized type (Gould 1996; Lewontin 2001). In individual-centered epidemiologic studies, the (average) risk in exposed individuals relative to the (average) risk in non-exposed individuals (relative-risk) is supposed to adequately represent the etiologic force of the exposure for each and all individuals of that population.
The Multifactorial Model of Disease Causation

In addition to aging, overeating, under-exercising, and cigarette smoking were considered candidate causes of the rise in degenerative diseases (Morris 1964). To measure the effects of these hazards, variables were chosen to represent exposure to them, either directly (number of cigarettes smoked) or indirectly (serum cholesterol levels to represent fat intake, and blood pressure to represent salt intake and/or stress). Epidemiologic studies conducted during the 1950s and 1960s revealed clusters of these variables, more or less strongly associated with each other and with CHD. According to Stallones (1980b), MacMahon’s “web-of-causes” represented the only possible way of modeling the epidemiologic findings, once there was no biologic structure to guide the way that the interacting variables should be ordered. According to this multifactorial model of causality, CHD was conceived as resulting from possible alternative and interconnected chains of exposures (component causes) that would eventually become sufficient to initiate the disease in an individual—conceived in Platonic terms as the average individual representative of the whole (Rothman 1982; Rothman and Greenland 2005). The multifactorial model of CHD has survived to today, with “new” risk factors incorporated into the “web-of-causes” with the “old” ones (Fruchart et al. 2004).

According to Kuller (1987): “the concept of multifactorial etiology of many chronic diseases may be a measure of our ignorance of causality rather than a biological principle” (p. 363 [abstract]). The fact that, 40 years after its introduction, we still lack a sound biologic theory to explain the development of CHD and its progression to death, allows us to question both this causal model and the degenerative idea behind it.

A New Framework for Disease Causation

In opposition to the Platonic understanding of nature, during the past years we have seen the emergence of a Darwinian perspective, which proposes that reality is made of variation and not of average, “typical” individuals, and that development and survival entails interactions between nature and nurture (Gould 1996; Levins and Lewontin 1985; Lewontin 2001; Meaney 2001; Tauber 1997). These two concepts may revolutionize the way we think about disease causation. An example from the area of cardiovascular diseases was recently provided by Rosen (2002): by accepting the idea of variability, and conceiving of diseases as products of interactions between specific substrates and triggers, atrial fibrillation and ventricular tachycardia were reclassified from two mere electrocardiographic patterns into several distinct diseases. Likewise, evidence in favor of conceptualizing CHD as several diseases, associated with specific vulnerabilities and triggers with varying distributions over time and space, will be presented below.
The decline in CHD mortality was accompanied by changes in important features of the disease. Hypercholesterolemia, endothelial lipid infiltration and plaque size—as well as the idea of CHD being a degenerative disease—were replaced by high triglycerides/low HDL, endothelial inflammation, and plaque instability in myocardial infarction (MI) patients (Libby, Ridker, and Maseri 2002). Sudden deaths fell more dramatically than did incidence of acute MI or long-term post-MI deaths, but prevalence of previous MI also decreased, continuing improvement in survivorship was documented, and the age of death progressively increased (Elveback 1979; Kattainen et al. 2006; McKinley, Mckindley, and Beaglehole 1989; Pepine 1997; Rosamund et al. 1998).

Interestingly, changes mirroring the ones described above occurred during the rise in CHD mortality. Before it was considered a degenerative condition, the “atheromatous affections of arteries” had been described in 1858 as inflammatory by Rudolf Virchow:

one who knows that the fatty degeneration is here only a termination, and that the process is really a formative one, inasmuch as it begins with a proliferation—he can readily imagine the possibility of another termination, namely ossification. As soon as the real ossification exists, we cannot help regarding the process as one which has arisen out of irritation of the parts stimulating them to new, formative acts; so far therefore it comes under our ideas of inflammation, or at least of those processes which are extremely nearly allied to inflammation. (p. 25)

As stated by Osler in 1908: “four great factors [were involved] in the causation of atherosclerosis—the normal wear and tear of life, the acute infections, the intoxications [including smoking, diabetes mellitus, obesity], and those combinations of circumstances which keep the blood tension high” (p. 430)—the same factors that are blamed for CHD today.

Spain (1960) compared CHD cases identified during autopsies done in two periods, 20 years apart. During the rise in CHD mortality (1931–1935), the odds of finding hypertensive disease among cases of coronary occlusion was 5:1. In 1951–1955, the degree of atherosclerosis in males from all age categories was considerably greater, but the odds of finding hypertensive disease among cases had reversed to 1:2.5. As Spain (1960) commented: “Since there is no reason for assuming that the incidence of hypertensive disease had significantly decreased over this period, we are perhaps justified in concluding that this findings indicate a considerable increase in coronary atherosclerotic heart disease with occlusion that is entirely independent of hypertension as a factor in its pathogenesis.”

During the 20th century, then, the profile of the CHD cases varied from the pre-epidemic period (until the 1920s and 1930s) to the height of the epidemic...
period (1950s and 1960s), and again during the decline in CHD mortality (1980s to date). Cases described during the pre- and post-epidemic periods were mostly associated with “inflammation” and an individual phenotype suggestive of insulin-resistance (obesity, diabetes), hypertension, and/or smoking. During the epidemic period, hypercholesterolemia became the hallmark of CHD cases, plaques showed higher degrees of occlusion, and case-fatality increased.

How should we interpret trends in CHD occurrence and the accompanying changes in features of CHD cases during the 20th century?

**Interpreting the 20th-Century CHD Trends**

Within the Platonic framework, and according with the multifactorial model of disease causation, changes in CHD occurrence and in features of CHD cases should be interpreted as corresponding to temporal changes in the availability of exposures (component causes) capable of interacting to produce new CHD cases among similar individuals.

The Darwinian view, on the other hand, attributes disease occurrence to interactions between exposures and dissimilar individuals, they themselves selected and modified by previous environmental exposures. According to Tauber (1997): “in a post-Darwinian world . . . there is no stable norm, no given, no static entity. . . . In each instant the organism must encounter its environment and itself to evolve, that is, respond. Life is dialogical—an ongoing and ceaseless dialogue between organism and its world” (p. 288). The Darwinian approach increases significantly the space for variability within the CHD category, and it strongly suggests that cases arising in individuals which are themselves a priori different possibly represent different diseases.

Diversity among atherosclerosis cases has been suggested before. According to Metchnikoff (1910):

> Probably diseases of the arteries of different kinds, and arising from different causes, are grouped under the terms atheroma and sclerosis. In some cases the lesions are inflammatory and are due to the poisons of the microbes. An example of such an origin is the case of the syphilitic sclerosis, in which the specific microbes (spirilla of Schaudinn) lead to precocious senescence. In other cases the arteries show phenomena of degeneration resulting in the formation of calcareous platelets which interfere with the circulation of the blood. (p. 31)

Following Metchnikoff’s lead, if instead of a multifactorial condition we conceived of “CHD” as encompassing different diseases associated with vulnerabilities and triggers with varying distributions over time and space, how would we model the rise and fall in CHD mortality and the changes in the main features associated with CHD cases during the 20th century?

Gould (1996) suggested that, within a Darwinian framework, a trend should be seen not as an entity moving somewhere, but as the result of relative expan-
sions and retractions of different sub-populations. We suggest that changes in CHD features accompanying the changes in CHD occurrence resulted from expansions and retractions of sub-populations of CHD cases over time. The steep rise in CHD mortality seen after the 1920s resulted from a huge expansion of a particular subpopulation of cases, characterized by high cholesterol levels and high case fatality (Azambuja 2004).

**Individual Vulnerability to Chronic Diseases**

*Identity begins with the genome, but the self arises through individual history and experience.*

—Tauber (1994)

A Darwinian model of disease causality presupposes interactions between environmental exposures and individuals with varied vulnerability to those exposures. While individual vulnerability is routinely considered in modeling infectious diseases, it is frequently ignored with respect to chronic diseases. Vulnerability to chronic “degenerative” diseases is usually conceived as a genetic attribute and, as such, not subject to short-term variation. According to Rose (1985): “There is a broad tendency for genetic factors to dominate individual susceptibility [to CHD]” (p. 34). Stehbens (1993) also equated vulnerability to genetics: “genetics, not dietary fat, appears to be the principal determinant of individual serum cholesterol levels in free-living humans” (p. 1342).

If one believes that vulnerability to CHD is a genetic characteristic, then the only explanation for variations in CHD mortality would be equivalent variations in some causal exposure. If vulnerability to hypercholesterolemia is genetic, than the decrease in serum cholesterol levels accompanying the decline in CHD mortality rates could only be attributed to dietary change (Carroll et al. 2005; Jousilahti et al. 1996). Evidence, however, does not support either of these claims. According to Stallones (1980a): “hypertension does not fit the trend of the mortality from ischemic heart disease at all; physical activity fits only the rising curve; smoking fits both (but is not in accord with the equivalence in the decline in mortality for the genders); serum cholesterol fits the falling curve” (p. 49). However, the decline in average serum cholesterol levels happened regardless of significant changes in dietary habits (Taubes 2001).

If variations in exposure alone can not explain the rise and fall in CHD mortality, and the Darwinian model is to be applied to chronic diseases causation, vulnerability cannot be conceived as a genetic attribute only.

Over the past years, different pathways have been postulated that could change individuals phenotypically and favor vulnerability to chronic diseases and CHD, including autoimmunity upon infection, metabolic reprogramming upon infection, and metabolic reprogramming associated with nongenomic mechanisms of inheritance (Azambuja 1998, 2004; Barker 2005; Barker et al. 1989; Drake and Walker 2004; Evans 1982; Fabricant et al. 1981; Gurevich, Pleskov, and

An alternative explanation of the huge expansion of a specific subpopulation of CHD cases occurring from the 1920s to the 1960s could depend not as much on changes in exposure to CHD “risk factors” as on an acquired window of vulnerability to the disease opened in the middle of the 20th century. If that were the case, what would its cause be? And how would it be related to the high serum cholesterol levels and fatality characteristics of those cases?

**Lipid versus Thrombosis Hypothesis for CHD in the 1950s and 1960s**

Notwithstanding its presumed multifactorial etiology, “the growing research endeavor on CHD that started after the World War II centered almost exclusively on the atheroma and, because of its lipid content, on the place of dietary fat and blood cholesterol levels” (Meade 1992, p. 287). The “diet-heart” hypothesis postulated that “hard, saturated fat elevates blood cholesterol. This, in turn, raises the likelihood that cholesterol would clog arteries, a condition known as atherosclerosis, which then increases risk of coronary artery disease, heart attacks and ultimately, death” (Taubes 2001, p. 36). (More sophisticated versions of the lipid hypothesis would be advanced during the 1980s [Steinberg and Witzum 1990], after the first publications of Ross and Glomset’s (1976) “response-to-injury” hypothesis and the identification of a “scavenger” receptor on macrophages, involved in the internalization of modified LDL [Goldstein et al. 1979].)

While the diet-heart hypothesis has remained the dominant hypothesis of CHD causation during the whole period, in the 1950s British investigators had already found evidence suggesting that atherosclerosis involved several related, though independent, processes. According to Morris (1964): “the main distinction would be into (1) lumen-occlusive lesions, which lead often to ischemic heart disease, and (2) mural atheroma which matters less” (p. 143). Morris believed that atheroma on the walls of the coronary arteries was far commoner in the population than occlusive disease, but that the higher rates of CHD mortality in England from 1954 to 1956 was related to an increase in acute coronary thrombosis and chronic coronary occlusion (secondary to the organization of previous intravascular coronary thrombosis), independently of the size of the mural atheroma.

Considering the rise and fall of CHD mortality as an acute expansion and retraction of a subpopulation of CHD cases characterized by hypercholesterolemia and elevated rates of lethality, would it be possible, instead of choosing between the lipid and the thrombosis hypothesis, to reconcile findings like hypercholesterolemia and increase in mural atheroma documented after 1935, and mortality associated with increase in acute coronary thrombosis and chronic coronary occlusion independently of the size of the atheroma (Morris 1964; Spain 1960)?
A Hypothesis That Reconciles Lipids and Thrombosis

A cohort correlation between the 1918 influenza and 1920–1985 CHD mortality in the United States led one of us to propose that the 1918 influenza pandemic triggered a window of vulnerability to CHD among survivors of a particularly vulnerable subpopulation: people 15 to 44 years of age in 1918 who, for still unknown reasons, had the highest mortality rate in the 1918-19 influenza epidemic (Azambuja 1998, 2004; Azambuja and Duncan 2002; Reinert-Azambuja 1994). Pleskov, Bannikov, and Zaitzev (1994) have described similarities between the amino acid sequences involved in cell attachment of the viral hemagglutinin in some strains of influenza virus and those of apolipoprotein B involved in the LDL binding to high-affinity LDL receptors. We suggest that, in this particular subgroup of influenza cases, an expansion of specific subsets of B and T cells targeted to viral epitopes shared by the apoB-LDL receptor interface might explain the features of the later increase in CHD.

Accumulation and oxidation of apoB lipoprotein at sub-endothelial sites of coronary arteries is referred to as the initial process of coronary atherosclerosis in high serum cholesterol CHD cases (Steinberg and Witzum 1990). And autoimmunity to oxidized LDL has been associated with progression of atherosclerotic lesions (Gurevich, Pleskov, and Levaya 2005; Wick et al. 1995). Accordingly, we postulate that cross-reactive antibodies directed against the apoB-LDL receptor interface would lead to sub-endothelial co-accumulation of lipids and immune products, and result in lipid peroxidation. This sequence of events has been demonstrated in cases of Heymann nephritis, an autoimmune disease where the main autoimmune target—megalin/gp330—is also a member of the LDL receptor family (Kerjaschki et al. 1997).

Thus, in vulnerable individuals, cross-reactive antibodies (and T cells) elicited by the 1918 influenza infection and diverted to the apoB-LDL receptor interface might explain the elevated serum cholesterol levels (aggravated upon high-fat diets or reinfections) and lipid deposits co-localized at sites of viral penetration: the left side of the heart, coronary arteries and the aortic arch. But would that be enough to explain the high levels of (sudden, unexpected) deaths among the cases?

We propose that the high rates of CHD lethality seen during the 1950s and 1960s resulted from autoimmune-mediated endothelial inflammation and thrombosis, associated with influenza reinfections. The rise in CHD mortality in Massachusetts steepened after the 1918 pandemic (Figure 1). The proportion of the excess deaths occurring during influenza epidemics attributed to organic heart diseases grew from 1.6% in 1918–1919 to 18.4% in 1920–1929, and to 51% in 1957–1960 influenza epidemics (Collins 1932; Eickoff, Sherman, and Serfling 1961). From 1957 to 1966 there were seven influenza epidemics in the United States. Arteriosclerotic heart disease was the only specific condition,
other than pneumonia and influenza, showing significant excess mortality during these seven periods (Houseworth and Langmuir 1974). Excess cardiovascular deaths among persons aged 65 and older were also documented during the 1968–1969 and 1972–1973 influenza epidemics (Alling, Blackwelder, and Stuart-Harris 1981). During the period of decreasing CHD mortality, myocardial infarctions occurring in association with influenza epidemics became progressively concentrated among the oldest individuals of the population (Sheth et al. 1999).

Figure 1 shows that CHD and influenza and pneumonia mortalities had divergent and complementary trends from 1918 to the 1960s, suggesting that deaths attributed to CHD might belong to the influenza pool. The reduction in death rates from CHD might be due in part to the continuing decline in influenza activity and to the absence of extensive influenza epidemics after 1968 in the United States (Gordon and Thom 1975).

To reconcile the lipid and the thrombosis hypothesis, then, we postulate that both hypercholesterolemia and coronary thrombosis reported among the CHD cases registered from the 1920s to the 1960s were secondary to an autoimmune response to influenza (re)infection developed among a specific subgroup of vulnerable individuals (those 15 to 45 years old in 1918) marked by the 1918 influenza pandemic virus. This would explain why at least two studies done during the 1960s found that the last or latest cholesterol measurement distinguished
CHD risk (and particularly the risk of CHD deaths) better than a high value for the average of all cholesterol measurements, “contrarily to the intuitive expectation” (Kahn and Dawbner 1966, p. 616; see also Keys et al. 1963).

**Lessons from the 20th-Century CHD Epidemic**

*In science, just as in art and in life, only that which is true to culture is true to nature.*

—Fleck (1935)

Challenging the powerful thought collective that attributed the rise in CHD mortality to a transition from what Armstrong, Conn, and Pinner (1999) call “an ‘age of pestilence and famine’ . . . to an ‘age of degenerative and man-made diseases’” (p. 62) would have been impossible before the 1980s.

Describing the British experience, Campbell (1963) reported that “CHD mortality rates doubled from 1890–1900 to 1927, doubled again by 1929, again by 1933, again by 1939, again by 1948 and again, by the sixth time by 1956, and continued [to rise] until 1959” (p. 535). While acknowledging that such change “is not very common in medical or biological statistics” and that it “may be found in early stages of the spread of an infectious disease among population with little immunity and, for a time, in the growth of a small population of animals with rapid reproduction rate and unlimited food supply,” Campbell concludes that “coronary heart disease had little resemblance to either of these conditions” and by attributing the trend to the spread of knowledge within the medical circle” (p. 535).

In spite of its steepness, the rise in CHD mortality was never really seen as an epidemic. Two things needed to happen so that the evidence could be reinterpreted: the downturn in the CHD mortality curve, which progressively disclosed its epidemic pattern; and the emergence of AIDS during the 1980s, which challenged the view of the “end of infectious diseases era” prevalent in the 1960s and 1970s and reopened the infectious/immunopathogenic path of inquiry to those studying chronic and coronary heart disease causation (Azambuja 2004; Capron 1993; Cunningham and Pasternack 1988; De Cross and Marshall 1993; Evans 1982; Javier Nieto 1998; Kuller 1987; Lorber 1996).

Study of the CHD epidemic provided information to characterize one of the subtypes of CHD cases: that associated with hypercholesterolemia and thrombosis, proposed here to be secondary to infection and autoimmunity. Without an event with the magnitude of the 1918 influenza pandemic, there would be no correspondent great expansion in the pool of CHD cases localized in time, and the disclosure of a process of infection and autoimmunity leading both ways—to hypercholesterolemia and to thrombosis—would probably be impossible.

Influenza might not be the only infectious agent associated with CHD. A group of relatively common pathogens (herpes, influenza, and papilloma viruses) were shown to share epitopes with the myelin basic protein; these viruses may
independently initiate and maintain the autoimmune process that results in multiple sclerosis (Wucherpfennig and Strominger 1995). Likewise, other common infectious agents (like herpes viruses, cytomegalovirus, and Chlamydia pneumonia) could, upon the right type of antigen presentation, might also contribute to the subpopulation of infectious–autoimmune CHD cases (De Geest, Van Linthout, and Collen 2003; Metha, Saldeen, and Rand 1998). Indeed, 22 infectious agents in addition to the influenza virus have been suspected of being implicated in atherosclerosis (Madjid et al. 2004).

If our hypothesis is correct, hypercholesterolemia might be a marker of autoimmunity in current CHD cases. This possibility deserves to be investigated, because it would allow us to identify and treat the subgroup of CHD cases with infectious etiology.

Smoking (and vulnerability to it) is known to have contributed to the pool of the mid-century CHD cases, and it may have been responsible for the increasing male/female ratio of CHD mortality observed during the whole period. Higher rates of respiratory diseases among smokers may have mediated the association between smoking and CHD. But other biologic mechanisms should be investigated. Furthermore, hypertension has been associated to CHD cases during the entire period, but hemorrhagic stroke, a condition also highly associated with hypertension showed a continuing downward trend during the 20th century (Lawlor et al. 2002). This suggests a fall in its contribution to the pool of CHD cases too, which might be attributed to a fall in the population vulnerability to it.

One subgroup of CHD cases, however, has grown since the 1980s: that associated with “inflammation” and insulin resistance. The growth of this subtype of CHD cases does not seem to be merely relative or secondary to the impressive fall in the number of hypercholesterolemic cases. The increase seems real, and if a lesson is to be learned from the 20th-century CHD epidemic, it would be that this increase should also be investigated as an epidemic. We need to ask what caused the emergence of the vulnerable (insulin-resistant) phenotype underlying the rising rates of obesity, diabetes, and their associated pool of CHD cases in the population. An important contribution may come from Barker’s hypothesis and from other theories of nongenetic inheritance of vulnerability across generations (Barker 2005; Barker et al. 1989; Drake and Walker 2004). Evidence suggests that the interaction of genes and the pre- and postnatal environments, against a heterogeneous background of vulnerability, could operate to produce metabolic programming effects across generations and result in obesity, diabetes, and cardiovascular diseases later in life. But one still needs to explain the inception of the rising rates and their simultaneous worldwide occurrence.

1 “Inflammation” is written in quotes to indicate that it must be considered provisional until we have learned more about the etiology and physiopathology of this specific type of CHD.
Coronary Heart Disease (CHD)—One or Several Diseases?

References


