

Scientific Disagreement and Evidential Pluralism: Lessons from the Studies on Hypercholesterolemia

Veli-Pekka Parkkinen[†]
v.k.parkkinen@kent.ac.uk

Federica Russo[‡]
f.russo@uva.nl

Christian Wallmann[†]
c.wallmann-520@kent.ac.uk

ABSTRACT

Inconsistencies between scientific theories have been studied, by and large, from the perspective of paraconsistent logic. This approach considered the formal properties of theories and the structure of inferences one can legitimately draw from theories. However, inconsistencies can be also analysed from the perspective of modelling practices, in particular how modelling practices may lead scientists to form opinions and attitudes that are different, but not necessarily inconsistent (from a logical point of view). In such cases, it is preferable to talk about disagreement, rather than inconsistency. Disagreement may originate in, or concern, a number of epistemic, socio-political or psychological factors. In this paper, we offer an account of the ‘loci and reasons’ for disagreement at different stages of the scientific process. We then present a controversial episode in the health sciences: the studies on hypercholesterolemia. The causes and effects of high levels of cholesterol in blood have been long and hotly debated, to the point of deserving the name of ‘cholesterol wars’; the debate, to be sure, isn’t settled yet. In this contribution, we focus on some selected loci and reasons for disagreement that occurred between 1920 and 1994 in the studies on hypercholesterolemia. We hope that our analysis of ‘loci and reasons’ for disagreement may shed light on the cholesterol wars, and possibly on other episodes of scientific disagreement.

[†] Centre for Reasoning, University of Kent, United Kingdom.

[‡] Department of Philosophy, University of Amsterdam, Holland.

1. Introduction

This paper investigates ‘scientific disagreement’, namely situations in which scientists hold different opinions or theories, or in which they hold different views about scientific results, or even about the processes to obtain such results. Disagreement may often, but need not always, lead to mutually inconsistent opinions concerning various hypotheses. To exemplify our theoretical claims about scientific disagreement, we use the recent debate on evidential pluralism and, more specifically, we focus on the long-standing controversy in medicine on the causes and effect of hypercholesterolemia, also known as the ‘cholesterol wars’.

There are several reasons that motivate investigations of disagreement between scientists. On the one hand, this extends a closely related debate in philosophy of science concerning inconsistent theories and what one can legitimately infer from those. The literature on inconsistencies in science has been focusing on what happens to a theory (most typically, a theory in physics) when some of its assumptions contradict each other. Specifically, what consequences does this entail? By and large, scholars in the field worked towards a logic of inconsistency. Focusing on inconsistent theories, rather than on disagreeing opinions or theories of certain scientists, misses out a host of interesting questions for philosophers of science. For instance, instead of investigating the root causes of advocating inconsistent theories, it considers only the end product, i.e. the theory. Also, such narrow focus may miss a subtle point: different opinions may not necessarily be incompatible with each other, but may lead to incompatible conclusions or results. Also, disagreement may come in degrees, which is a broader issue than the ‘binary’ consistent vs inconsistent one. On the other hand, recent debates on evidence in medicine have focused on agreement, that is, what is needed in order to establish a causal claim: typically, scientists use multiple sources of evidence, which are, at least in principle, of equal importance. In a sense, evidential pluralism has been occupied with a question about consensus, or how a community comes to agree on what causes what. Hence, focusing on what may hinder the scientific community from establishing such claims broadens the debate on evidential pluralism. Let us develop further.

Consider the debate on inconsistencies in science first. The typical problem addressed in this literature is the inconsistencies of hypotheses of physical theories, including cosmology or chemistry. Most often the question has been

about whether theories with inconsistent sub-theories should be outright rejected or, if not, in what way they may undermine our realist intuitions (see e.g. Gregersen and Køppe (1988); Brown (1990, 2002); Norton (2002)). So, while traditional discussions indeed aimed at finding a logic for inconsistent theories (Meheus, 2003), some other scholars tried to frame the debate differently, thereby broadening the focus. For instance, Smith (1988) tries to shift the focus to the practice of experimental science in order to understand why inconsistencies may occur. To give another example, Nersessian (2002) distinguishes two ways of looking at inconsistency: the perspective of logic, where inconsistency is seen largely detrimental in the reasoning process, and the perspective of history of scientific development, where inconsistency can play important heuristic roles in the process of conceptual change.

Smith (1988) and Nersessian (2002), we submit, mark a considerable broadening of the narrow focus on the logic of inconsistencies. For one thing, it is an approach more in line with the ‘History and Philosophy of Science’ and the ‘Philosophy of Science in Practice’ schools. An important consequence of this is that potential inconsistencies are not just properties of theories, but may instead be related to modelling practices, or to the way modellers phrase a research question or set up an experiment, or to the complex socio-political dynamics that take place in scientific practice. This shift of focus in the ‘inconsistency debate’ from logic to scientific practice offers us an entry point, provided that one broadens the scope: what is at stake is not merely the inconsistent, logical, status of theories.

The traditional debate on inconsistencies in science mainly focused on ‘abstracta’: the formal properties of theories and the structure of inferences we draw from theories. This formulation of the problem of inconsistencies focuses on theories as the main, or even sole, interest of philosophers of science (see e.g. Boon (2015)) and largely ignores scientific practices and agents, which may be the root causes for the observed inconsistencies at the level of theories. What we are interested in, instead, is how scientists come to form opinions and attitudes that may, but need not necessarily (or logically) contradict each other. These may concern particular hypotheses or causal claims. That is, we aim to investigate the factors that explain why disagreement within a scientific community occurs and how it is sustained.

Contributions such as those of Smith or of Nersessian are pivotal in that they make a first step in broadening the scope of the debate on inconsistencies from logic (of inconsistent theories) to the scientific practice. Such move must be

taken a step further. So the suggestion is to change, from the start, the focus of the problem: instead of talking about inconsistencies, we talk about disagreement among scientists or within the scientific community.

One goal of this contribution is to explain where and why disagreement may occur in the course of the scientific process. We are especially interested in cases where scientists consider evidence coming from many different types of sources: How do scientists evaluate multifarious evidence in *establishing* causal claims in such cases? How do differences in opinion about the weight of different types of evidence give rise to disagreement over causal conclusions? And how is the weighing of evidence justified? Thus, our analysis ties to the ongoing discussion and debate concerning evidential pluralism, specifically in the health sciences.

Consider now the current debate on evidential pluralism, which, as a matter of fact, has been focusing on agreement. The classic discussion of Russo and Williamson (2007) explicitly tackled the question of what evidence should support a causal claim, in order to consider it as established or confirmed. Subsequent contributions, e.g. Clarke et al. (2014), tried to make this position more sophisticated by elaborating on the very notion of evidence or how different sources of evidence (notably, evidence of difference making and of mechanisms) mutually help each other. Likewise, critical authors such as Broadbent (2011) or Howick (2011) attacked the argument about the need of both dimensions of the evidence, but still focused on *establishing* a causal claim. Admittedly, questions about disagreement have not been tackled specifically by either side of the debate (supporters and critics of evidential pluralism).

Possibly, the only case study analysed in the debate about evidential pluralism where questions about disagreement arose is the famous ‘Simmelweis case’ (Thagard, 1998; Gillies, 2005; Russo and Williamson, 2007; Broadbent, 2011). Simply put, the story is that Ignaz Semmelweis, a doctor active in nineteenth-century Vienna, had hypothesised that puerperal fever was due to some kind of infection. Even though he couldn’t support his proposal with a mechanism compatible with the medical knowledge available at that time, he suggested washing hands appropriately after performing autopsies and before assisting women in labour as a precautionary measure. The scientific community at that time, however, precisely because a solid theoretical framework to support Semmelweis’ hypothesis and his precautionary measure was lacking, resolved to reject his ideas. By and large, the philosophical debate centered on the question whether or not the scientific community was right in

rejecting Semmelweis. However, one should arguably dig deeper and try to understand what, exactly, the community had divergent views on. Such questions may concern several aspects: why does disagreement arise? But also: what should we do in such cases? Later in section 2 we shall discuss several reasons (besides evidence) that may hinder consensus about causal claims. In the subsequent sections, dealing with the ‘cholesterol wars’, we shall also note, when appropriate, how cases of disagreement may be resolved.

As announced earlier, within the debate on evidential pluralism, we specifically focus on the studies on hypercholesterolemia, i.e. high levels of cholesterol in the blood, and its role in heart disease. Hypercholesterolemia has causes and effects that have been difficult to establish; some, to be sure, are still under debate. Donald Gillies (2011) discusses the same episode in the context of evidential pluralism. His main point is that, to reach consensus, the scientific community only needed a plausible rather than a well-confirmed mechanism. While we agree with that, we also think that important questions remain about disagreement in this case. This episode of medical research is therefore instructive in two ways: to get a fair theoretical understanding of ‘disagreement’ with the help of historical reconstruction of scientific controversies, and to anticipate ways of handling disagreement in present-day controversies. In what follows, we use ‘the cholesterol hypothesis’ to refer to the somewhat imprecise package of ideas, hypotheses, and studies about the relations between diet, serum cholesterol, and heart disease. Simply put, according to the ‘cholesterol hypothesis’, hypercholesterolemia is a cause of heart disease, and, quite possibly, high blood cholesterol level causally depends on diet.

Our reconstruction of the history of the hypothesis that hypercholesterolemia causes heart disease uses as a main thread the question of how to establish a causal claim and, via the discussion of ‘loci and reasons’ of disagreement, aims at highlighting what hindered consensus. While the scientific community now agrees on a number of features of this medical condition, we hope to make clear that we are in no way offering a Whiggish reconstruction of the controversy. Also, our discussion focuses mainly on epistemic factors. Yet, in the history of science and philosophy of science, disagreement has often been attributed to non-rational or extra-scientific factors. Among the most prominent non-rational factors reside incompetence, ideology and venality (Mumpower and Stewart, 1996). Among the most prominent extra-scientific factors reside moral, religious, political or metaphysical consequences of the theory or claim in question Lugg (1978).

Our choice is not guided by questions of importance or relevance. ‘Pure’ epistemic factors are, to some extent, easier to access for us philosophers. But we also hope that, if we get the epistemic discussion about ‘loci and reasons’ right, then it will be also easier to discuss non-epistemic factors, in future work. In particular, in this case, questions about the interference between interests of pharmaceutical or food industries and public health concerns should get proper attention and dedicated discussion.

The paper is organised as follows. In section 2, we present the ‘loci and reasons’ for disagreement, namely where and why scientists may disagree. In section 3, we offer a historical reconstruction of the ‘cholesterol wars’, emphasising crucial points in the long search for the causes and effects of this medical condition. In section 4, we combine the theoretical discussion of the ‘loci and reasons’ and the historical reconstruction of the cholesterol wars by illustrating some of the places and reasons of disagreement in the scientific community. The discussion here is not meant to be exhaustive, but hopefully provides a first step in how to use our theoretical framework. Finally, in section 5, we summarise the main arguments and identify further lines for research.

2. Scientific disagreement: loci and reasons

In this section, we develop a nuanced picture of the loci and reasons – where and why – scientists may disagree. From the start, we abandon terms such as ‘inconsistent’ or ‘inconsistency’ and instead couch our analysis in terms of how, where, and to what extent scientists (dis)agree on a given hypothesis, claim, process, or output. In sections 3 and 4, we shall focus on disagreement that is not attributable to non-rational or extra-scientific factors, which are typically due, amongst others, to sociological and psychological factors. For instance, it is not uncommon that public or private authorities and metaphysical background theories influence the positions advocated by scientists.

To analyse scientific disagreement, we distinguish two different but interrelated questions:

- (i) At which stages of the scientific process do scientists disagree?
- (ii) Why do scientists disagree?

To answer question (i), we reconstruct scientific process as a form of step-wise problem-solving process. To answer question (ii), one may recur to epistemic,

sociological, psychological, or other factors.

We discuss these two questions in this very order. We try to keep loci and its reasons separate. Section 2.1 discusses loci of disagreement and section 2.2 discusses reasons for disagreement and their relations to the loci. It is worth clarifying from the start that loci and reasons for disagreement are clearly not independent of each other, but rather interconnected. And even within these categories, items are not independent of each other. This section therefore runs an exercise in conceptual analysis so that the intersections become more visible when needed, namely in the historical reconstruction of the hypercholesterolemia case in section 3.

2.1 Loci of disagreement

To begin with, we reconstruct the scientific process in most general terms, as finding a solution to a specific problem. This consists of the following steps:

1. Define the problem.
2. Search for relevant data and clues to solve the problem.
3. Evaluate the data and clues.
4. Draw conclusions from the body of evidence.

This characterisation of the scientific method is admittedly very general. It is on purpose and it has a virtue: it is widely applicable across scientific domains, from physics to medicine. While in line with a Popperian reconstruction of scientific method, and with subsequent hypothetico-deductive accounts, our reconstruction does not reduce to a strict deductive approach. Thus, ‘drawing conclusions’ is not only a matter of logical deduction, as in a Popperian account, and can instead involve a number of inferential practices, including, say, analogical reasoning or inference to the best explanation. Furthermore, from a strict Popperian perspective, the rejection of a given hypothesis (as formulated at the ‘problem definition’ stage) would lead to an outright rejection of the whole theory. Instead, according to the characterization above, this does not automatically happen. As we shall explain further in this section and later in the discussion of hypercholesterolemia, disagreement may happen at various stages of the scientific process and for various reasons. Also, the rejection of one hypothesis (or piece of evidence) does not lead to the outright rejection of all available knowledge about the phenomenon under scrutiny.

We thus take a broad characterisation of the scientific process as a useful starting point to develop on the issue of (dis)agreement; specifically, in section 4, we bestow attention to how disagreement concerned issues related to evidence or to disease mechanisms. In so doing, we already depart from the standard set up of the debate on inconsistencies: our interest is not in presumed or actual logical incompatibilities of theories, but rather in critical places of the scientific process, where disagreement is likely to occur.

Our reconstruction of the scientific process is closely related to that of psychologist and policy analyst Thomas Stewart (1991) but also differs in some important respects. Stewart proposes the following hierarchy for scientific judgment regarding global warming:

- Level 1: raw data and facts (where, according to him, no disagreement happens yet).
- Level 2: studies and results are grouped.
- Level 3: interpretation and aggregation of particular lines of research.
- Level 4: drawing broad conclusions based on lower level conclusions.
- Level 5: policy recommendation based on Level 4 conclusions are reached.

The first four levels of the hierarchy presuppose a clear problem definition and raw data and facts that are not subject to disagreement. However, as we shall see in section 3 and 4, this did not happen in the cholesterol controversy. In order to account for some of the substantial disagreement in the cholesterol controversy, we cannot assume that problem definitions and Level 1 are free from disagreement. Instead, our contribution is to show that loci and reasons need to be made explicit, as important disagreement may arise at those stages too. Below, we discuss the main loci where disagreement is likely to occur and preview some loci of historical disagreement in the cholesterol case. The controversies on the ‘cholesterol hypothesis’ help us illustrate this point. This lends further support to our general strategy to move away from the narrow focus on inconsistent theories and instead broaden the perspective by taking the whole scientific process, and all the stakeholders involved, into account.

Philosopher Andrew Lugg (1978) discusses the issue of scientific disagreement. In this paper, Lugg presents a classical view of disagreement, according to which rational researchers disagree only if they differ with respect to the data they possess, i.e., only if they disagree with respect to the locus “search for relevant data and clues to solve the problem”. He then goes on to reject the classical view by considering three historical cases for which he argues that 1) the same data was available to the

researchers, 2) different disciplinary backgrounds were the key reason for major disagreement, and 3) disagreement rooted in different disciplinary backgrounds was rational in these cases. Our aim, at this stage, is largely descriptive, rather than normative. This means that we do not discuss whether this classical view is correct, i.e., whether it is rational or irrational to disagree, especially with respect to the third and fourth step of problem solving. In other words, we do not take issue with the question whether, given a certain problem and a certain body of data and clues, there is only one rational way to evaluate the data and clues to answering the problem. We now proceed to the discussion of steps 1–4 of the scientific process, as potential loci for disagreement.

Problem definition. In the first step, different epistemic agents may consider different problems, because they use relevant concepts differently or they consider different questions altogether (Mumpower and Stewart, 1996). Consider, for instance, the question whether high cholesterol causes heart disease. Depending on how the concept of cause is understood, the same research question may point to different problems. We will see below in section 3 that part of the historical disagreement about the cholesterol hypothesis was due to different meaning and use of the word ‘cause’.

Search for relevant data and clues to solve the problem. In the second step, scientists may have access to different data and clues. This may be due to different reasons. One reason is that a group of researchers may not be aware of the existence of, or may not be able to understand, some data or clues which are relevant to the problem. For instance, it might be that due to their disciplinary focus and training, epidemiologists are less inclined to attend to bio-chemical data about the mechanism connecting cholesterol to heart disease than cardiologists are. A second reason is that, a group of researchers may have different methodological standards as to how we can create evidence relevant to the problem or which evidence they consider to be relevant to the problem. Are, for instance, quasi-experimental, observational methods, animal studies, mechanistic studies or qualitative methods suitable to create data or clues relevant to causal questions? We will see below in section 3, that some researchers did not accept the cholesterol hypothesis, because, although plenty of data from observational evidence was available at their time of evaluation, no trial evidence was available.

Evaluate the data and clues. In the third step, methodological questions about quality of evidence become pertinent. If clues and data are considered to be relevant to the problem, the question remains whether they are of good or bad quality. When evaluating the data, scientists may assign, for instance, higher quality to data

obtained by trials than to evidence from biological mechanisms or observational studies. Apart from assigning different qualities to different types of evidence, researchers may disagree whether a certain piece of evidence is of high quality or low quality. While in the cholesterol case some researchers judged the Lipid Research Clinics Coronary Primary Prevention Trial to be of high quality, opponents of the cholesterol hypothesis pointed to its possible flaws - see also later section 3.

Draw conclusions from the body of evidence. The fourth step of problem solving concerns evidence amalgamation, i.e., the question how evidence from different sources should be combined to solve the problem. This problem is hard-wired. For instance, currently the most widely used proposal for evidence amalgamation proposed by the GRADE-working group and the methods employed by the International Agency for Research on Cancer (IARC) still differ with respect to their use of observational data and data from mechanistic studies to solve problems. Daniel Steinberg, a pioneering investigator in the field of lipid metabolism and atherosclerosis, argues that, in the cholesterol case, all available evidence should be considered (Steinberg, 2007, p.1-3). Data from animal studies, observational studies, mechanistic studies and trials has been available since long time. However, as we will see in section 3, there was disagreement as to how to combine these different sources of evidence. It is important to note that disagreement about amalgamation of the evidence need not necessarily concern the whole body of evidence and different types of evidence. For instance, some researchers claimed that results from early trials were mixed, while others claimed that the same trials convincingly lend support to the cholesterol hypothesis - for a fuller discussion, see also section 4.

At each of these steps, reasons for disagreement are multiple. As anticipated in the introduction, some reasons for disagreement are due to epistemic factors, others to socio-political factors (including vested interests of individuals or of lobbies), or psychological factors. In sections 3 and 4 we mainly focus on epistemic factors, which include different training of the scientists, different methodological views, different conceptualisations of cause, mechanism, or other. We now proceed to analyse in detail possible reasons for disagreement and the steps of the problem-solving process that they most affect.

2.2 Reasons for disagreement

We borrow an initial list of reasons for disagreement from the works of Lugg (1978) and Mumpower and Stewart (1996):

- Epistemic;
- Disciplinary differences;
- Fact-value confusion;
- Sociological;
- Psychological;
- Individual (in)competence;
- Extra-scientific.

Lugg groups these reasons into non-rational (sociological, fact-value confusion, psychological and individual incompetence), extra-scientific, and rational (certain disciplinary differences and epistemic). In what follows we do not make any normative claim about which of the reasons is a rational reason for disagreement and which not. We use this list of reasons in the discussion of the controversy on the ‘cholesterol hypothesis’ that will follow in section 4, focusing specifically on epistemic, disciplinary differences, and fact-value confusion.

Sociological (and political) reasons for disagreement are reasons that are attributable to the socio-political structure and dynamic of the scientific environment or the broader community the researchers are part of. For instance, authorities may dominate the field of the researcher in question or economic interests of the food industry may interfere with dietary recommendations. Psychological reasons concern researchers’ individual minds.

Individuals may disagree because they think differently about the problem, without being incompetent, self-interested or advocating personal values (Mumpower and Stewart, 1996). Specifically, researchers may have different modes of cognition, i.e., they may be creatively or analytically orientated minds or more or less subject to general biases in human reasoning like confirmation bias (Mumpower and Stewart, 1996)¹.

While researchers are held to be less responsible for their mode of cognition or their susceptibility to confirmation bias, individual incompetence may range from logical fallacies to misreading and misinterpreting different clues. There may be, of course, other kinds of social or psychological constraints at stake.

Lugg (1978) identifies the following extra-scientific reasons: different metaphysical, religious, moral or political background theories. Lugg considers, for

¹ Mumpower and Stuart speak of different organizing principles rather than psychological differences.

instance, different training and different background assumptions and theories within different disciplines as most important source for disagreement. Researchers from different disciplines often “have different access to the system of scientific belief and practice as a whole” (Lugg, 1978, p.282).

To argue for the claim that different disciplinary background is a key reason for disagreement, he considers three examples. In particular, Lugg considers landscape modelling. As discussed by Lugg, Agassiz’s view that the glacial action extensively modelled the landscape (for Agassiz’s view see Rudwick, 1969) has been rejected by Lyell in favour of his iceberg theory (for more information about the relevant views see Davies (1969)). Lugg argues that this disagreement is due to the fact that Agassiz was an expert in paleoichthyology and Lyell in topography (Lugg, 1978). Different training of researchers may lead to disagreement at all four steps of problem solving. As we will see below in sections 3 and 4, cardiologists and epidemiologists, for instance, are interested in different questions, they employ a different concept of cause, they refer to different studies and evaluate them differently.

Fact-value confusions occur if researchers mix up what should be the case and what is the case. A classical example is a conflict of interests. Researchers may benefit from the acceptance of a claim by, for instance, their peers or a wider public. This may consciously influence the researchers’ judgment (for instance, in case of funding by a third party) or unconsciously cloud the judgment of the researcher (for instance, in case of the reluctance to admit misjudgment). Often, fact-value confusions may lead different researchers to ask different questions; also, the extent to which problem solving is influenced by fact-value confusions remains highly controversial, argues Mumpower and Stewart (1996). For instance, in the cholesterol controversy, researchers are often concerned with side-effects of cholesterol reduction. While this is a natural concern for whether reducing cholesterol is good or bad for a certain given population (a value-question), it is irrelevant as to whether cholesterol causes heart disease (a fact-question). For instance, substantial disagreement on whether high cholesterol increases mortality by other causes than heart disease remained until the first statin trials (see section 3). Whether high cholesterol increases mortality by other causes that are unrelated to heart disease is clearly irrelevant as to whether heart disease is caused by high cholesterol.

For the sake of the argument, we assume that if two researchers, belonging to the same discipline, share the relevant socio-political environment and psychological features, are not subject to fact-value confusion or incompetence and no extra-

scientific factors are pertinent disagree, they disagree for purely epistemic reasons. But our argument is not to show that such (genuine) epistemic disagreement is possible. It is likely that epistemic disagreement is ultimately explainable in terms of non-epistemic reasons. But we pragmatically classify disagreement as epistemic, if there is no reason to assume that one of the other reasons can account for the disagreement. In this way, we aim to show that even remaining at the epistemic level, a proper account of disagreement must take into account the whole scientific process, rather than just focusing on the final product, i.e., the theory, as typical of the literature on inconsistencies in science. This should lend further support to our initial choice of shifting the focus from inconsistency to disagreement.

3. The “cholesterol wars”

Contemporary biomedical science spends considerable effort on studying the causes and effective treatments of heart disease. Today, a consensus view states that heart disease can be effectively treated or prevented altogether by controlling blood cholesterol levels, especially cholesterol carried in low density lipoproteins (LDL), for example by statin therapy. However, historically, the question of what causes heart disease has neither been obvious nor unambiguous. Until the early decades of the 20th century, the prevailing view stated that the thickening of arteries and the resulting symptoms of heart disease are an inevitable consequence of the loss of elasticity of blood vessels that occurs throughout the vascular system as one ages. In other words, ‘heart disease’ was not a disease category per se, in the sense of a pathological phenomenon whose specific aetiology needs to be uncovered in order to treat and prevent it; it was instead an irreversible side-effect of old age. First attempts to reconceptualise heart disease as a preventable, disease-like phenomenon were motivated by animal experiments suggesting that dietary factors can greatly accelerate the development of arterial lesions that are the likely cause of symptoms like infarction and stroke. We describe these and further developments below.

3.1 Evidence for the “cholesterol hypothesis”: the early 20th century

The earliest evidence for a causal link between cholesterol and heart disease came from animal experiments performed by Russian pathologist Nikolai Anitschkow. Anitschkow conducted a series of experiments in which rabbits were kept on a high cholesterol diet after which he studied their arteries in autopsy for changes in

the structure of the vessel wall (Anitschkow, 1913).

The experiments showed that cholesterol-fed rabbits had developed lesions in the arteries reminiscent of human atherosclerosis – a common cause of heart disease. Studying the composition of the lesions, Anitschkow discovered deposits of free cholesterol, as well as accumulation of macrophages laden with cholesterol and other lipids, leading him to conclude that cholesterol probably causes the lesion development. However, a detailed explanation of the mode of cholesterol deposition or the role of the macrophage activity was beyond the medical understanding of the time.

Anitschkow's results flew in the face of the received view among his contemporaries, according to which the thickening and hardening of arteries that characterizes atherosclerosis was due to the loss of elasticity of blood vessels and various 'wear and tear' that inevitably occur as one ages. The prevailing 'senescence hypothesis' was founded on population-wide statistics and clinical observation: incidence of heart disease and the loss of elasticity of blood vessels was robustly correlated with age. Unsurprisingly, Anitschkow's experiments attracted a fair amount of criticism. It was argued that a rabbit – a herbivore whose natural diet is cholesterol-free – is hardly a valid model for studying the effects of cholesterol in humans, and that the experimental intervention was so unrealistically severe that little could be said about cholesterol's effects in the range that would mirror average human diet (Ophuls, 1933; Weiss and Minot, 1933). Further animal experiments were subsequently conducted by many scientists on different model species, partly addressing these worries (Aschof, 1933; Bruger and Oppenheim, 1951). These showed mixed results, but many accorded with Anitschkow's original experiments. However, due to the lack of understanding of cholesterol metabolism and the mechanism of cholesterol transport in the blood, the animal data was open to many skeptical interpretations (Duff and McMillan, 1951; Peters and Van Slyke, 1946).

While caution concerning an extrapolation from the animal models was justified, the successful experiments could be further validated by comparison to findings about heart disease in humans with familial hypercholesterolemia. Familial hypercholesterolemia is a near-perfectly heritable condition involving extremely high innate blood cholesterol level and xanthomas, i.e. visible accumulation of cholesterol and other lipids under the skin. Familial hypercholesterolemia is often accompanied with atherosclerotic heart disease at a very young age. This was known, and was clearly consistent with the idea that heart disease has at least something to do with cholesterol metabolism (Muller, 1939). Plenty of room for disagreement nonetheless remained: with no mechanism identified, the association between the

xanthomas, blood cholesterol and heart disease was open to many causal interpretations, and one would perhaps not pick as first choice an interpretation that contradicted the received view. In addition, the validity of these findings was unclear, also considering the role of cholesterol in nonhypercholesterolemic people (Steinberg, 2007, p. 31). These findings - the observational data of hypercholesterolemia patients and the animal experiments - were the most important early evidence that founded the cholesterol hypothesis as a contender to the prevailing senescence hypothesis, even if the hypothesis initially had but few proponents. In section 4, we relate some aspects of the early history of the cholesterol hypothesis to the taxonomy of scientific disagreement introduced in section 2.

3.2 In search of mechanisms and the first trials

In the early part of the 20th century, virtually nothing was known about the mechanism of cholesterol transport in the blood. This would not directly undermine the hypothesis founded in Anitschkow's work, but it rendered further research on the cholesterol-heart disease link somewhat speculative. This, along with the well-entrenched view that arterial thickening is due to ageing, might explain why the cholesterol hypothesis was not investigated in a clinical setting until much later. The opportunity costs of setting up large clinical studies are high, so some initial credence is required of any hypothesis that deserves to be tested in a clinical trial.

Insights into the cholesterol transport mechanism came with the discovery of low-density lipoproteins by biologist John Gofman and his collaborators, who subsequently described a taxonomy of plasma lipoproteins according to their density (Gofman et al., 1949). Gofman was already convinced that cholesterol plays a causal role in atherogenesis, and proceeded to correlate different lipoproteins with heart disease outcomes, the main finding being that heart disease risk is elevated in patients in which cholesterol is predominantly carried in mid- to low-density lipoproteins. Once these findings became available, the cholesterol hypothesis gained the initial theoretical credibility to justify larger scale clinical research.

Clinical trials testing the hypothesis in humans were not conducted until some forty years after Anitschkow's initial studies, but the time between the 1950's and 1970's saw several trials testing the ideas that reduction of dietary fat intake or switching unsaturated for saturated fats would reduce the incidence of heart

disease (see Connor and Connor (2002)). We will restrict our discussion to a basic outline. A complete survey of all these studies should be the object of a separate historical investigation.

These trials mostly sampled either populations of hospital patients, or individuals who had survived previous cardiac events (and who presumably thus were motivated to comply with the dietary intervention); individuals in the trials were administered a low fat diet or diet rich in unsaturated fats. Blood cholesterol, clinical cardiac events and/or mortality were typically measured as the outcome. The patients in the intervention group either served as their own control group, or in some cases of inpatient trials, were compared to a control group who were given regular hospital diet. The results from these trials were mixed: many of the studies showed results right above the threshold of statistical significance. This, at face value, somewhat undermined the prospects of the cholesterol hypothesis as a basis of clinical decisions or public health policy.

While the evidence from diet intervention trials was mixed, there were some large scale epidemiological studies conducted roughly at the same period that seemed to corroborate the cholesterol hypothesis. One of the most important of these was the ‘Seven Countries Study’ lead by Ancel Keys (1966). This study compared populations from seven countries with markedly different dietary traditions with respect to saturated fat intake, and calculated the correlation between blood cholesterol and incidence of heart disease. The study showed that heart disease mortality was roughly proportional to average blood cholesterol levels as well as saturated fat intake, a fact that was clearly explainable by the cholesterol hypothesis. However, due to its observational nature, it was difficult to rule out, beyond any doubt, the possibility of confounding causal factors. Keys at al.’s study measured several other dietary, health, and lifestyle factors, but as the theoretical understanding of the disease process was far from complete at the time, it was difficult to tell which confounders one would have to be able to control in order to gain strong evidence of causation. The most obvious confounder - genetic background - could nonetheless be later ruled out by pooling the original Keys study together with evidence gleaned from populations that share the same genetic background but differ with respect to dietary habits. One example of such a design was a study by Robertson et al., which compared the incidence of heart disease in native Japanese populations and migrant Japanese population residing in California (Robertson et al., 1977). Aside from the work by Keys’ group, perhaps the other most important epidemiological study was the Framingham Heart Study, a longitudinal study conducted by the National Heart Institute in Framingham, Massachusetts (Kannel et al.,

1961). In the first cohort the study included more than five thousand subjects from Framingham, who were measured for several potential risk factors, one of them being cholesterol, and followed these with periodic examinations for more than twenty years. The results that emerged showed a clear association with baseline cholesterol levels and subsequent myocardial infarction.

At this point it is important to distinguish two separate questions about cholesterol's role in the aetiology of heart disease:

1. Does serum cholesterol cause the damage to the arteries?
This question concerns the causal power of cholesterol to induce - one way or another - the lesions seen in experimental animal and human atherosclerosis.
2. Does serum cholesterol level (strongly) depend on dietary fat intake?
This question concerns the influence of diet on blood cholesterol levels.

Before the development of effective cholesterol-lowering drugs, researchers were limited to dietary interventions for lowering blood cholesterol, and thus any hope of answering question 1 with clinical trials depended on a positive answer to question 2. This is a fairly simple point, but ignoring it can create confusion. Even if studies employing dietary interventions failed to show reduction in symptoms or mortality from heart disease, this would not necessarily mean that the cholesterol hypothesis is false, if it is taken as an answer to the first question. Also, even if one considers the problem to concern question 2, short lasting follow up trials may not be able to provide unequivocal answers, as it was somewhat unclear how long one must adhere to a low fat diet in order for it to show beneficial effects. During the time when the mechanism of cholesterol-induced heart disease was yet largely unknown, it was easy to lump these questions together, and thus equivocate between two problem definitions that each ask for an answer to just one of the questions. This could lead to situations where evidence pertinent to just one of the problems was brought in to evaluate the other, or both problems. As a result of this, incompatible opinions concerning whether high cholesterol causes heart disease that are not content-related may result. As we have explained in section 2, tacit differences in the problem definition may sustain scientific disagreement.

Despite considerable uncertainty about the status of the cholesterol hypothesis, the American Heart Association (AHA) cautiously acknowledged it in 1961, stating that elevated blood cholesterol should be considered a clinically relevant factor. This

conclusion was based on the animal data and the results from epidemiological studies. However, AHA did not consider it to be conclusively shown that heart disease is preventable by dietary measures. Page et al. (1961, p.133) concerning a diet with poly-unsaturated fats state:

This recommendation is based on the best scientific information available at the present time. More complete information must be obtained before final conclusions can be reached.

In 1971, the National Heart, Lung and Blood Institute launched an interventional study titled The Lipid Research Clinics Coronary Prevention Trial (LRC-CPPT) - named after the participating centers - in which hypercholesterolemic men were treated with a cholesterol-lowering agent and followed for more than seven years. The compound used in the intervention condition was cholestyramine, which hastens the removal of cholesterol in bile acid. It was hypothesized that this would lead to overall reduction in serum cholesterol, thus providing a way to bypass the diet-blood link and testing the causal effect of serum cholesterol as such. The results showed a reduction of circa 20 per cent in the primary outcome of death, or nonfatal myocardial infarction (Rifkind, 1984). This could be considered a more definitive demonstration, but the total evidence still provided mixed signals, as each source of evidence had its characteristic limitations, and individual studies provided somewhat inconsistent results. These are summarized in table 1. As evidence from any particular source could be contested by pointing to their characteristic limitations, the disagreement would persist.

A prominent cardiologist, John McMichael, expressed a highly skeptical take on the evidence in the *British Medical Journal* at the end of the 1970's:

The best-conducted dietary trials under the auspices of the MRC's statistical control have given convincingly negative results. In survivors who have had coronary manifestations, and are thus at special risk, low-fat and soy-bean oil diets, which can lower the blood cholesterol concentration, have been entirely ineffective in slowing the progress of the disease towards recurrences or death. Drugs that reduce the blood cholesterol concentration also failed to influence outcome when tried on large numbers of similarly affected patients whose disorder was liable to deteriorate (McMichael, 1979, p. 173).

Table 1

<i>Source of evidence</i>	<i>Limitations</i>
Animal models	Questionable validity
Observational studies	Confounding by factors unrelated to lipid intake/ cholesterol
Clinical trials	Direct interventions on blood cholesterol not feasible; non-significant effects

To back up his case, McMichael cites several trials conducted by The Coronary Drug Project, sponsored by the National Heart, Lung and Blood institute. These trials tested the effects of interventions with estrogen, niacin, clofibrate, and dextrothyroxine sodium on total mortality and specific coronary heart disease (CHD) mortality. Of these studies, only the niacine and clofibrate trials were carried out as planned. Neither showed statistically significant effects with respect to the designated endpoints (The Coronary Drug Project, 1975). It should be noted that at the time of publication of McMichael's critique, a similar trial with cholestyramine – mentioned above – was underway, and would in time produce much more promising results. It is unclear whether data from the cholestyramine trial was publically available yet at that time, and thus we must withhold judgement on whether McMichael deliberately chose to ignore those data or not.

What is clear however, is that McMichael was well aware of the epidemiological evidence correlating cholesterol, saturated fats and CHD. In other publications, he went on to claim that the endorsement of the cholesterol hypothesis by various public health bodies was largely due to misattribution of evidential weight to epidemiological data in the absence of corroborating theoretical explanation or definite demonstration of causation in a controlled trial (McMichael, 1979, p. 174). McMichael's stance towards the whole body of evidence could be reconstructed as one where statistical evidence from epidemiology could not on its own support conclusions about policy. Instead, according to McMichael, statistical evidence ought to be interpreted in light of other evidence – a process that McMichael somewhat vaguely calls 'scientific analysis' – before any firm conclusions about causation can be drawn. This scientific analysis, according to McMichael, had been neglected by those who accepted the cholesterol hypothesis on the basis of (mostly) just the epidemiological evidence (McMichael, 1979, p. 174).

We take McMichael's reasoning to exemplify one possible model for evaluating evidence of causality, in which a relatively high degree of coherence of observational evidence, trial evidence and theoretical understanding of the underlying mechanisms is required before one is allowed to infer a causal claim. This is perhaps also characteristic of his disciplinary background of cardiology, where a pressing concern is to understand the details of the processes that lead to disease outcomes in the individual. For epidemiologists or public health experts - traditionally much more preoccupied with identifying risk factors based on population-wide studies - such requirements were less salient, which could explain a coarse distinction between the supporters of the cholesterol hypothesis in the epidemiology community on the one hand, and a group of staunch critics within cardiology on the other hand.

McMichael followed this attack to the evidence base of the cholesterol hypothesis by questioning the safety of the recommended cholesterol-lowering diets (McMichael, 1979, p. 174). This was not an appeal to an alleged epistemic shortcoming of the cholesterol hypothesis, and thus gives the impression that McMichael was asking for extra-scientific values to bear on a scientific question, amounting to the kind of fact-value confusion we described in section 2. However, such arguments are not uncommon in medicine, nor should they be, as medical science aims not only at understanding disease phenomena, but also at using that understanding to reduce the burden of disease. Thus, one can find in the literature two types of arguments against the cholesterol hypothesis, often not kept apart: one contests the truth of the hypothesis that hypercholesterolemia causes heart disease, and the other contests the clinical implication of this hypothesis - the normative claim that one should intervene to lower blood cholesterol. Conflating these two questions is apt to lead to confusion over what problem, exactly, the debaters disagree on: on the status of the causal claim, or on what follows from the claim in terms of implications for good clinical practice.

3.3 The bio-chemical basis of the cholesterol hypothesis

The next breakthroughs came from mechanistic studies in biology. In 1972, Joseph Goldstein and Michael Brown discovered the LDL receptor (Goldstein and Brown, 2009). The LDL receptor is a cell-surface receptor that allows cells to extract cholesterol from LDL, to be used for cell wall synthesis. Based on what was known of cholesterol biosynthesis, Goldstein and Brown had assumed, and then demonstrated, that endogenous cholesterol synthesis in humans is

feedback-regulated. That is, there is a mechanism by which excess cholesterol is extracted from the bloodstream, and endogenous cholesterol synthesis adjusted with respect to the amount of cholesterol that is so extracted. Identifying the LDL receptor as a component that allows cells to scavenge cholesterol from blood presented Goldstein and Brown with an explanation for the heritability of familial hypercholesterolemia: a mutated LDL receptor gene would leave the liver cells of its carrier without the capacity to extract cholesterol from blood, and endogenous cholesterol biosynthesis would thus be unable to adjust to exogenous cholesterol intake, leading to a continuous increase in overall blood cholesterol. This would also confirm the direction of causation in the epidemiological studies on familial hypercholesterolemia - the subcutaneous cholesterol accumulation must be caused by excess blood cholesterol, not the other way around.

The early 1970's brought another pivotal discovery in atherosclerosis research, one that curiously did not directly connect with the cholesterol hypothesis. This was the discovery of the platelet-derived growth factor, which suggested a mechanism for the development of the mature atherosclerotic lesion that is characterized by abnormal smooth muscle cell proliferation. Russell Ross and John Glomset hypothesized that the formation of atherosclerotic lesions begins with an injury to the cell wall. Platelets recruited to heal the injury subsequently release signaling molecules, such as the platelet-derived growth factor (PDGF), which induces smooth muscle cell proliferation and thus the growth of the lesion (Ross and Glomset, 1976a,b). Ross and Glomset's hypothesis is based on a general mechanism of cell signalling known to control growth and cell differentiation. Ross and Glomset's innovation was to apply this mechanism schema to the question of smooth muscle cell proliferation. This hypothesis was then backed up (among other evidence) by the discovery of the crucial entity PDGF whose characteristic activities can explain the growth of the atherosclerotic lesion in particular. Ross and Glomset hardly mention lipoproteins in their research, merely pointing that sustained high levels of serum lipoproteins might be among many other factors that dispose to the initial damage.

Ross and Glomset's hypothesis became very influential among cardiologists, who were traditionally more skeptical of the cholesterol hypothesis. It is instructive to pause for a moment to consider this, and compare the situation in epidemiology, in which the cholesterol hypothesis gained the support of authority figures like Ancel Keys from much earlier on. Cardiology's job

description is to study the causes of specific clinical cardiac events such as infarction and stroke, and thus its focus in atherosclerosis research has been on the properties of the advanced lesion which are directly responsible for the timing and severity of such events. Important research questions here concern the process of smooth muscle cell proliferation and the deposition of connective tissue to form the lesion that leads to clogging of the artery, and the propensity of rupturing of the fibrous cap that covers it. Ross and Glomset brought an existing mechanism template from cell biology – cell-to-cell communication via signaling molecules – and applied it to the case of the developing lesion in a way that immediately suggested an explanation for the clinical events cardiologists were interested in. Even though lipoproteins were acknowledged as one possible factor among many that might initiate the injury, in Ross and Glomset's hypothesis these factors were conceptualized as background conditions that enable the mechanism of lesion development to operate, not as salient causes of the clinical events.

Compared to cardiologists, epidemiologists were on average more likely to accept the cholesterol hypothesis. Epidemiologists consider as their task the mapping of health outcomes to risk factors, and the primary research heuristic is to look for statistical dependencies that predict incidence of disease, or to devise mathematical models of the spread of a disease at the level of whole populations. Epidemiologists' research efforts are not primarily organized around elaborating mechanisms and applying established mechanism schemas to new phenomena, and the research is thus not directed by the availability of applicable mechanism schemas.

From these considerations, one can put together possible explanations for the difference in opinion between the two fields when it comes to acceptance of the cholesterol hypothesis. Cardiologists were, qua disciplinary framework, more focused on proximate causes of heart disease, and required the articulation of a mechanism before an explanation could be accepted. The proximate cause of symptoms of atherosclerotic heart disease is the fibrous cap that forms on top of the initial fatty streak lesion. Cardiologists' main interest was in the properties and development of the latter, which is the proximal cause of symptoms of heart disease. By conceptualizing atherosclerosis as a phenomenon crucially involving cell-signaling, Ross and Glomset were able to explain many properties of the mature lesion, as well as its development, in terms of mechanisms that were already widely accepted, and which involved no essential appeal to the role of cholesterol. By contrast, epidemiologists were

focused on tracking more distal factors which make a difference to the distribution of CHD in different populations, and could plausibly accept cholesterol as such a factor even in the absence of detailed mechanistic understanding. One could thus see the difference in opinion between the two groups as a difference in problem definition, and difference in focus on proximal versus distal causes. Later in section 4 we shall discuss these differences in more detail.

Further mechanistic studies provided evidence of ways by which cholesterol carried in low density lipoproteins (LDLs) could induce the formation of lesions in the artery wall. One major line of research focused on the 'oxidative modification hypothesis of atherogenesis'. The oxidative modification hypothesis suggested that while cholesterol itself is chemically inert in a way that could not cause the initial damage to arteries, oxidized forms of LDL would be able to penetrate the vessel wall and start the process of development of the complex lesion. Studies dating in the late 1970s, and from there on, demonstrated that LDL gets oxidized *in vivo* in model organisms, and that oxidized LDL is both capable of damaging endothelial cells and being taken up by macrophages *in vitro* (Chisolm and Steinberg, 2000).

These studies could further bridge the gap between already known mechanisms of atherogenesis and the large-scale epidemiological results. If excess cholesterol in itself could set in motion the mechanisms of lesion formation, this would give a straightforward explanation of the epidemiological data in terms of a concrete causal process. But as with all laboratory research, the validity of the *in vitro* results with respect to whole-organism physiology, and that of the animal models with respect to humans, was somewhat uncertain.

3.4 Contemporary consensus and the statin era

By the early 1980s, the cholesterol hypothesis could be supported by multimodal evidence from laboratory studies, some human experiments, and epidemiological studies. While the hypothesis was not understood as being conclusively demonstrated, public health authorities clearly acknowledged the literature supporting it. In 1982, the Nutrition Committee of the American Heart Association (AHA) issued a statement on AHA's current stance concerning the relationship between diet and heart disease (Grundy

et al., 1982). This statement lists elevated blood cholesterol, high blood pressure, diabetes, and obesity as the main CHD risk factors, notes the special role of LDL in atherosclerosis, as well as the effects of saturated fat intake on blood cholesterol (Grundy et al., 1982, p. 16-17).

The road to large clinical trials capable of compelling a consensus on the cholesterol hypothesis required the development of effective drugs for lowering serum cholesterol levels, i.e. the discovery and clinical development of statins. Statins are a class of drugs that inhibit endogenous cholesterol biosynthesis by competitively binding HMG-CoA (3-hydroxy-3-methylglutaryl-coenzyme A). HMG-CoA is a precursor of mevalonic acid, which again is a precursor of cholesterol in the cholesterol biosynthesis pathway.

Normally, the conversion of HMG-CoA into mevalonic acid is catalyzed by another enzyme, HMG-CoA reductase (HMGR). Statins mimic the structure of HMGR, but do not have the same enzymatic activity. Administering statins reduces the rate of cholesterol production, as HMGR is a rate-limiting enzyme in the pathway. The first statin, nowadays known as compactin, was initially discovered by biochemist Akira Endo while working at the Sankyo Research Laboratories in 1972 (Endo, 2010). Its development for clinical use was somewhat delayed, possibly due to unpromising results in preclinical animal testing (Endo, 2010, pp. 487-488). Once the compound was tested in human trials, the results looked immediately promising (Mabuchi et al., 1981; Yamamoto et al., 1980). These early studies were small, but positive enough to warrant more extensive studies. Possibly, the most important single study compelling a consensus on the cholesterol hypothesis was the Scandinavian Simvastatin Survival Study (4s) (Scandinavian Simvastatin Survival Study Group, 1994). This was a randomized trial with simvastatin - one of the 21 later statin drugs - involving 4444 participants who were followed up for an average of 5.4 years. The study was a success, its major advantage being that it could demonstrate not just a reduction in mortality from heart disease or any proxy endpoints, but a reduction in overall mortality. This helped clear worries about the safety of cholesterol reducing interventions. Since then numerous trials have provided concordant results, persuading a majority of the medical community to accept the causal role of hypercholesterolemia in heart disease (Collaborators, 2005).

The statin studies cemented the opinion that aggressive reduction of

serum cholesterol reduces the incidence of heart disease due to coronary narrowing in many at-risk groups. As of today, no major disagreement exists within mainstream medicine concerning such a causal claim. Residual controversies about the efficacy of dietary regulation, however, have continued over the statin era. Nonetheless, the recent opinion of major regulatory and public health authorities generally favors a view according to which reduction of dietary cholesterol is effective as well. In 1990, the AHA stated that the evidence for a causal link from either dietary or serum cholesterol to CHD is overwhelming and justifies national programs for cholesterol control (LaRosa et al., 1990). The most recent AHA dietary recommendations still advice limiting saturated fats and trans fats (Eckel et al., 2013). However, trial evidence suggests that the beneficial effect of dietary cholesterol regulation is quite modest (Hooper et al., 2001).

4. Loci and reasons for disagreement in the cholesterol wars

The cholesterol wars did not come to an end, but some armistice happened. Slowly, the scientific community reached some consensus about aetiology, preventive measures, and interventions. Consensus, however, is not an unassailable fortress. Periodically, competing hypotheses get dusted off (e.g. ‘sugar conspiracy’, as also recently reported in *The Guardian*²) to question the very basis of this consensus. Therefore, it is useful to linger a bit more on these controversies and to refer back to the loci and reasons of disagreement. In this section, we consider more explicitly where certain institutions or researchers in the past have disagreed about the link between cholesterol and heart disease. Specifically, we show that substantial disagreement occurred at all loci of disagreement.

Looking retrospectively at the cholesterol wars, it seems to us that by 1994 the causal link between cholesterol and heart disease had been established with reasonable confidence - our stance here relies especially on the work of (Steinberg, 2007). However, we do not mean to imply that certain institutions or researchers were plainly wrong in doubting the causal link. On basis of the available and incomplete evidence, there was room for rational disagreement.

² See for instance: <https://www.theguardian.com/society/2016/apr/07/the-sugar-conspiracy-robert-lustig-john-yudkin>, accessed on 27 November 2016.

4.1 Different problem definitions

As section 3 has shown, the so-called ‘cholesterol hypothesis’ does not refer to a single, specified research question. Our reconstruction revealed instead the existence of rather different questions, for instance:

1. Is high cholesterol a cause of heart disease? (A causal question)
2. Will lowering cholesterol lower heart disease mortality? To what extent? (An interventional question)
3. Will lowering cholesterol lower overall mortality? (A safety question)
4. In which group of people will lowering cholesterol lower heart disease mortality? (An extrapolation question)
5. Does diet lower cholesterol? (A different causal question)

The first four questions are ordered according to the amount of evidence needed to answer them. Increasingly more evidence is needed to answer questions further down the enumeration. Many researchers considered answering the causal question to be the easiest task (see, for instance, Page et al. (1961) and US Department of Health and Human Services and others (1984)). According to the NHI consensus conference in 1984, the causal relationship between cholesterol and heart disease may hold even if it is not the case that lowering cholesterol levels lowers heart disease rate. We read in the main document:

Our conclusion that reduction of blood cholesterol levels will reduce the rate of coronary heart disease is based partly on the evidence for cause-and-effect presented above and partly on the direct evidence from clinical trials noted below. (US Department of Health and Human Services and others, 1984)

At first glance, an answer to the interventional question seems to require results from trials. As rightly pointed out by many opponents of the cholesterol hypothesis (McCormick and Skrabanek, 1988; Mitchell, 1984; McMichael, 1979), consistent results from trials were not available at their time (at least until the results of the Coronary Primary Prevention Trial were published in 1984, see section 3). In section 4.2, we discuss whether a positive answer to the interventional question requires results from trials in more detail.

Even if lowering cholesterol in turn lowers heart disease, it is not ipso facto

the case that mortality by causes other than heart disease risk is also lowered. At first glance, a definite answer to the question whether cholesterol lowering is free of serious side effects required evidence that wasn't available until 1994, when results from the 4s trial became available (see section 3).

The Coronary Primary Prevention Trial (see section 3) seems to have, for many, established that lowering cholesterol will lower heart mortality in men aged 35-39 with no history of coronary disease and no signs of current disease and that were already at high risk (total blood cholesterol level of 265 mg/dl or higher). It was by no means clear whether the trial in combination with the rest of the evidence licenses the extrapolation of the result to other groups like, for instance, to women.

Part of the historical disagreement on cholesterol may be due to the fact that researchers had different questions in mind. A researcher, for instance, rejecting the claim that the relationship between cholesterol and heart disease is causal may highlight that the effect of lowering cholesterol on lowering overall mortality has not been established (see, for instance, the reconstruction of McMichael's position in section 3). While this is certainly a valid point against recommending a population-wide cholesterol diet or treatment with cholesterol lowering drugs, it is not relevant to the causal question.

Part of the disagreement concerning the causal question above is due to different uses of the concept 'cause'. An important distinction exists between the notions of proximal and distal causation. As discussed in section 3, cardiologists were mainly concerned with proximate causes of disease: these causes would provide the details of the biological processes responsible for triggering symptoms of heart disease. By contrast, epidemiologists were concerned with finding risk factors such as environmental exposures that are associated with disease outcomes, regardless of the distance – temporal or mechanistic connection – between the exposure and the disease outcome.

For instance, Mitchell discusses many beliefs about cholesterol and heart disease he thinks would be shown to be false by the year 2000 (Mitchell, 1984). One of those is that CHD is caused by atherosclerosis. Mitchell claims that this is wrong because myocardial infarction and sudden death are rapid events. According to him, thrombosis is the relevant cause, as the occurrence of thrombosis is the close proximate cause of infarction and similarly rapidly developing events. To support the claim that atherosclerosis is in the relevant sense a cause of infarction, one would have to show how exactly atherosclerosis leads to sudden obstruction of blood flow similar enough to thrombosis – i.e.

one would have to establish the mechanism of plaque rupture.

From an epidemiological point of view - in which mechanistic details are often deliberately black-boxed - similar relevance criteria do not apply. Other authors seem to advocate a stronger concept of cause. They require, for instance, that a cause is the most relevant cause of its effect (or at least a cause associated with large effect-size) or that a cause leads deterministically to its effect. For instance, McCormick and Skrabanek seem to advocate a deterministic concept of cause. According to them, the effect must inevitably follow the cause:

Even infectious diseases do not inevitably follow exposure to pathogens, the necessary cause, because many other conditions have to be satisfied before disease becomes manifest [...]. Because these risk factors are simply associated with an altered probability for the disease and have not been shown to have a causal relation, the term should be dropped and replaced by “risk marker”. (McCormick and Skrabanek, 1988, p.839)

However, even if researchers had the same question in mind, substantial disagreement remained.

4.2 Inference from the evidence: Establishing a causal claim without trials?

At different stages of the discussion main researchers drew different conclusions from the current evidence. Basically, in the cholesterol case, not one single piece of evidence from a single source could be taken to be sufficient to establish a causal claim, or to license a reasonable expectation about the result of an intervention. Critics rightly point out that each source of evidence is subject to certain flaws (see section 3). The question about disagreement is therefore also a question of amalgamating evidence, i.e., merging evidence from different sources. Especially, conclusions obtained by consensus conferences were often based on combining evidence from different sources. In this section, we take a closer look at disagreement that concerned how to combine evidence from different sources.

Some opponents of the cholesterol hypothesis pointed out that establishing a causal claim requires results from interventional trials. For instance, McCormick and Skrabanek claim that

While epidemiological studies may lead to the formulation of important

hypotheses about the causes of coronary heart disease, only experiment can prove causal relations. (McCormick and Skrabanek, 1988, p. 839)

In a similar vein Mitchell claims:

Confronted with an association which could either be causal or a marker, the only way forward is to mount an interventional trial. Mitchell (1984, p. 292)

If consistent evidence from interventional trials is needed to establish causality, then Mitchell and McCormick seem to be right in rejecting the cholesterol hypothesis, as mentioned earlier in section 3. However, doubt has been raised about this methodological principle by advocates of evidential pluralism. Evidential pluralism is the view according to which causal claims, in medicine but also elsewhere, are typically established on the basis of evidence of difference-making and of mechanism (Russo and Williamson, 2007). Simply put, evidence of difference-making establishes appropriate correlations between the putative cause(s) *C* and the putative effect(s) *E*. Evidence of mechanism, instead, provides hints as to how *C* causes *E*. This thesis is epistemological in character and does not reduce causation to any of these evidential components, nor to their conjunction.

If the mechanism by which *C* is connected to *E* is sufficiently well established, this could, in connection with other types of studies, provide compelling reasons for a causal interpretation of the link between *C* and *E*. Specifically, in the cholesterol case, if we have sufficient evidence about cholesterol synthesis and metabolism, as well as its biochemical capacities, then, together with observational and animal studies we should have strong enough reasons for predicting the effect of an intervention that lowers cholesterol on heart disease. The issue at stake here is not that different sources of evidence point, independently, to the same causal relation. Instead, what is at stake is that we need to amalgamate different sources of evidence so that they together provide stronger support to a given causal claim. It is in this sense that evidential pluralists have been using the metaphor of ‘reinforced concrete’, a composite material whose resistance is due to the mutual support of steel and concrete, which resists different kinds of stress (Clarke et al., 2014).

As we allude to the amalgamation of different types of evidence as a way to arrive at more secure (causal) conclusions, our position bears important similarities to another prominent analysis of pluralism in science, namely,

integrative pluralism as advocated by philosopher Sandra Mitchell (2002, 2003). Before briefly presenting her view and further stressing points of contacts between our views, it is important to emphasise one important difference: our discussion considers pluralism with respect to sources of evidence, while hers is about pluralism of modelling frameworks and explanations. Mitchell's analysis starts from the fact that, within the life sciences, it is common to see a diversity of methods and models used to explain properties of a single target system. She then argues against a view according to which different models are in competition only if they are pitched at the same 'level of analysis', in the sense that they address the same explanation-seeking question. In Mitchell's example of the explanation of the division of labour in social insect colonies, the levels-of-analysis view would conclude that models of (division of labour in terms of) colony-level adaptation are not in competition with self-organization models of the same phenomenon. The former addresses a historical, phylogenetic question, while the latter answers a how-question about the ontogeny of individual colony organizations.

Mitchell then points out that this is not correct: as is the case with the diverse explanations of colony organization, an explanation pitched at a particular level of analysis may involve presuppositions that are incompatible with explanations at other levels, or may itself be incompatible with presuppositions made by explanations at other levels. In Mitchell's account, the question then arises how to account for the heterogeneity of modelling perspectives. Her answer is that such a plurality of models is expected in the study of a complex system – all scientific representations are partial descriptions of reality, and the very nature of complex systems requires that one employs a number of modelling perspectives, each of which focuses on a subset of the causal factors involved in generating the behaviour of a complex system.

The 'integration' part of integrative pluralism in her account stands for the following idea: given that different models are models of a single target, they are not independent of each other in the way that the levels-of-analysis framework suggests, but should instead ultimately be brought in line with each other in order to obtain a fuller understanding of the target – that is the integration. Mitchell's main focus is thus on the integration of diverse explanatory perspectives, not on the question of what sorts of evidence one needs in order to establish the validity of these explanations, i.e. the validity

of the causal ascriptions they make about the target phenomenon. By contrast, our interest lies in establishing causal claims at a specific level, e.g. the level of a population or of an individual, using evidence that is heterogeneous or multifarious. For one thing, this means that evidence is collected in different ways (e.g., by experimental or observational studies), or refers to different levels of analysis (e.g., by observing distributions of disease in a population or studying the progression of a disease in an individual patient), or comes from causally dissimilar sources such as animal models and minimal laboratory systems.

To return to evidential pluralism in our sense, important institutions that evaluated the evidence connecting heart disease with cholesterol in fact combined different kinds of evidence to assess the cholesterol hypothesis. For instance, in US Department of Health and Human Services and others (1984) the acceptance of the cholesterol hypothesis was based on the following evidence (see section 3):

- Genetic evidence: Investigating children with hypercholesterolemia;
- Animal model evidence;
- Epidemiological evidence: Framingham and others, new trials like
- Lipid research clinics coronary primary prevention trial;
- Mechanism of LDL discovered by Brown and Goldstein.

An opponent of the cholesterol-hypothesis may doubt the credibility of either piece of evidence (see section 3 for flaws associated with each type of evidence). But even if all of those studies could be discredited individually, it does not follow that one can discredit them in conjunction. Each of the studies provides evidence missing in other pieces of evidence, which is the core idea behind evidential pluralism. Indeed, Steinberg sees as main reason for the long lasting controversy:

Most important of all, resistance to the need to synthesize evidence of several different kinds – epidemiological evidence, experimental observations in animals, genetic evidence, clinical observations, and clinical trial data – in evaluating the true strength of the lipid hypothesis. The early clinical trial results while weaker than might have been desired, were nevertheless impressive if they were weighted in the context of all the other available lines of evidence. (Steinberg, 2007, p.197)

4.3 Evaluation of clues and evidence

In section 3, we briefly described various historically important examples from the overall evidence base for the cholesterol hypothesis, including experiments on model organisms, mechanistic studies, epidemiological studies, as well as clinical trials employing dietary or pharmacological interventions.

Different authors often disagreed on the correct interpretation of these studies in terms of how much weight should be given to different types of evidence when attempting to establish conclusions about causality. For instance, different authors had wildly varying views concerning the applicability of results from model organism research to humans. But aside from the disagreement about the interpretation of particular results, there were also controversies regarding the quality of the evidence. Crucially, these debates focused on trial evidence commonly assumed to be required for demonstrating causality beyond doubt. As some of these controversies concerned studies that have later become canonized as key pieces of evidence in favor of the cholesterol hypothesis, we mention some famous critiques below.

The first successful trial employing a (non-statin) pharmacological intervention was the Lipid Research Clinics Coronary Primary Prevention trial (LRCCPPT), mentioned in section 3, which treated hypercholesterolemic men with cholestyramine. The study has become one of the prime examples of a demonstration of the clinical benefits of cholesterol-lowering, while the most widely raised critical issue has to do with external validity. All subjects in the study were men and, almost certainly, genetically hypercholesterolemic, which raised concerns about the possibility of extrapolating the results to the wider population. Nonetheless, some commentaries raised significant critiques considering the quality of the study itself.

In an editorial to *Nutrition Today*, George V. Mann launched a harsh critique on the cholesterol hypothesis, singling out the cholestyramine trial as one key piece of faulty evidence (Mann, 1985). Mann accused the authors of the LRCCPPT of bad research practices, of choosing statistical techniques in light of what is most likely to produce a positive result, rather than according to a predetermined study design (Mann, 1985, p. 13). Mann's critique was based on an analysis by Kronmal (1985). Kronmal had pointed out that in describing the design of the trial, the authors state that a p-value below .01 in a one-tailed test is required to license the rejection of the null hypothesis, given the design

of the experiment. However, in a separate article reporting the results of the trial, the significance level used was .05. According to Kronmal's reanalysis of the data, the trial would not have reached a significant result in a .01 one-tailed test, or in a two-tailed test using the usual .05 cutoff for significance (Kronmal, 1985).

The authors of the trial also performed within-group analyses for both the cholestyramine treatment group and the placebo groups. These, too, showed a correlation between cholesterol levels and the measured CHD endpoints. Kronmal points out that some proportion of the correlation might be due to the way the authors had handled missing data points. Each of the test subjects was examined once in four months. If a subject failed to show up to two subsequent examinations, his baseline LDL level was used as the value of those month's measurements. Now, consider a situation in which a subject fails to show up to (many) examinations because they had experienced clinical events associated with CHD, and had failed to follow the treatment plan for this very reason. Kronmal writes:

If such an association between the precursors of the CHD event and withdrawal from therapy exists, then the results of the rule used would tend to make those who have a CHD event, as a group, have a lower mean reduction in LDL-C level than those who did not have a CHD event. (Kronmal, 1985, p. 2093)

A single example is of course not sufficient to represent the breadth of all the debate about the quality of evidence. We merely intend this example as an illustration that disagreement took place on two levels: one concerning the interpretation of results that were largely agreed upon as being well established, and the other concerning the quality of particular studies, i.e. whether or not a particular result can be taken as well established in the first place. These questions are often intertwined in practice, but it is nonetheless possible and analytically useful to distinguish them. A community of researchers might for example be in complete agreement about some result being well established in experimental animals or in a particular study population, yet disagree sharply about the relevance of the result for clinical applications.

4.4 Evidence of mechanism

We previously introduced the core idea of evidential pluralism. Central to this

is the use of evidence of mechanisms to establish causal claims. This may be a source of important disagreement. What is at stake here is not merely whether scientists genuinely disagree about the details of the mechanism leading from C to E - in our case, from cholesterol to heart disease. Instead, even when asking one single question (does cholesterol cause heart disease?) there may be different mechanisms at stake, and this is where disagreement about the evidence may lie.

To see how this is possible, one should note that the term ‘mechanism’ can be understood in two ways. On the one hand, by ‘mechanism’, we may mean a concrete causal process in the world: the totality of entities and activities, whatever they are, that somehow participate in bringing about some phenomenon. On the other hand, ‘mechanism’ may mean something like a diagram one finds in a molecular biology textbook: a rather abstract, truncated representation of some aspects of the total process. These theoretical descriptions of mechanisms are not designed to capture all the features of the concrete mechanism in the world. Rather, they represent information about particular aspects of the concrete mechanism, for the purposes of addressing particular explanatory tasks and for guiding further research. Our evidence of mechanisms comes from research programs organized around elaborating such theoretical mechanisms (Bechtel and Richardson, 2010).

When studying a complex phenomenon such as heart disease, there are often many theoretical mechanisms that apply to the phenomenon. Thus, there is no single research programme that we could consult to get evidence of the total, concrete mechanism that leads from the causes of heart disease to its clinical manifestation. Instead, we have patchy evidence of many theoretical mechanisms, operating at different levels of organization. In principle, one can reach different conclusions about the level of mechanistic support for a given causal claim by selectively attending to parts of this mosaic of mechanisms. This is interesting for our purposes, as possible disagreement here concerns evidential relevance – what kinds of evidence one should gather and consider, and what can be ignored – rather than the quality of the evidence, or how strongly it warrants causal inferences.

One might ask: which mechanisms should one attend to in order to answer the question does high cholesterol cause heart disease? There may be no unambiguous answer. The connection between hypercholesterolemia and heart disease is a complex phenomenon - the processes linking the cause to the effect involve operations at the levels of metabolism, immune system, vascular

system, and various gene regulation and cell signaling systems. Whole subfields of biomedical research are devoted to theorising about particular aspects of the global mechanism that mediates the dependency between cholesterol and heart disease. Research in each of these respective fields is organised around elaborating specific theoretical mechanism templates.

It is evident that the whole of research on the causes and effects of high cholesterol is not even considering the mechanism of heart disease in the sense of a concrete constellation of all the entities that are involved in the process. For example, as was mentioned in section 3, in the discussion of Ross and Glomset's endothelial injury theory, cardiologists' skepticism towards the cholesterol hypothesis might have been due to their particular disciplinary focus. This is very plausible given the structure and uses of mechanism concepts in science. What Ross and Glomset were able to do was to give a theoretical description of many properties of the advanced atherosclerotic lesion in a way that satisfies the strictures of mechanistic explanation. That cholesterol is downgraded as a background condition in Ross and Glomset's explanation is a consequence of applying a specific mechanism template of cell signaling: it does not follow that if Ross and Glomset are right, then cholesterol has only a minor role in the aetiology of heart disease.

The importance of cholesterol becomes evident once one recognizes that the early, clinically silent fatty streak is a background condition required for the development of the advanced lesion, and cholesterol has a crucial role in the development of the fatty streak. Our discussion of some of the loci and reasons for disagreement in the cholesterol wars – different problem definitions, evidence, evidence amalgamation, and evidence of mechanisms – should make clear that there isn't a 'unified' or 'coherent' theory of hypercholesterolemia. This, rather than being a special case, is quite common in medicine, and indeed across the sciences. Consequently, the move of shifting the focus from theories to modelling practices should liberate the traditional debate on inconsistency and open up a whole path of research that investigates disagreement, in its several dimensions.

5. Conclusion

In this paper we addressed the question of scientific disagreement, specifically the question of which factors may create and sustain a dissensus about a particular problem within a scientific community. On the one hand, our

discussion follows up the first step made in a certain strand in the literature on inconsistencies in science and broadens the debate by considering the role of modelling practices, and of modellers, and not just the logic of inconsistent theories. We made a firm choice in abandoning terms such as ‘inconsistency’, in favour of ‘(dis)agreement’. Disagreement is in fact broader in scope than inconsistency: opinions about hypotheses, theories, procedures, or results may diverge and yet not be (logically) inconsistent.

On the other hand, we consider this to be an important opportunity also to broaden the scope of the debate on evidential pluralism. Simply put, this is the view that to establish a causal claim one needs multiple sources of evidence. By and large, evidential pluralism has been dominated by the question of how to establish a causal claim, using multifarious evidence. In this paper, we tackled the question of where and why, in the scientific process, disagreement may arise, thus hindering the establishment of causal claims.

Section 2 provided a theoretical account of disagreement, including epistemic, socio-political, or psychological factors. We considered loci and reasons for disagreement at a very general level. We reconstructed the scientific process as a special case of problem solving and we argued that both loci and reasons of disagreement are manifold. In sections 3 and 4, we illustrated the ‘loci and reasons’ for disagreement using an episode from history and philosophy of the health sciences: the studies on hypercholesterolemia and its relation to heart disease. Our historical reconstruction shows that many of the loci and of reasons of disagreement discussed in section 2 played an important role in the controversies surrounding the relationship between cholesterol and heart disease. Indeed, after the first animal trials, it took nearly 80 years for the medical community to reach consensus that high blood cholesterol is a cause of heart disease. Section 3 retraces major breakthroughs in establishing that high cholesterol is a cause for heart disease. Among those were results from animal experiments, studies of familial hypercholesterolemia, observational studies, mechanistic evidence and trial results. Each of these sources of evidence, however, suffers from certain limitations. This fact contributed substantially to disagreement concerning the cholesterol-heart disease link.

For instance, part of the disagreement was due to the fact that there were at least five different questions guiding the research. Also, scientists were working with different concepts of cause or disagreed about the quality of the available data and about their interpretation. The major reason for the

longstanding controversy was, however, the reluctance of researchers to integrate evidence from different sources. This is an important point, given the current attention of the medical and philosophical communities to questions about evidence amalgamation.

Our account of the loci and reasons for disagreement and the discussion of the ‘cholesterol wars’ ultimately lend further support to evidential pluralism. Because a community may disagree for different reasons and at various stages of the scientific process, disagreement may be overcome by looking at all the available evidence, rather than at one piece at every one time. Our point, at this stage, is mainly epistemic.

It is important to note, however, that we did not try to answer why *de facto* leading scientists disagreed on the relationship of cholesterol and heart disease. In particular, we did not consider whether the course of history is best explained by epistemic or by pragmatic factors. We did, for instance, not aim to reconstruct the debate from a socio-psycho-political point of view. Instead, we mainly considered epistemic reasons for disagreement.

As explained in the Introduction, this is not meant to mark an order of importance or relevance. We agree that to get a full picture of disagreement in the ‘cholesterol wars’ (as well as in other cases) reasons for disagreement other than epistemic must be taken into account. For instance, one pressing question to address is whether leading scientists (especially opponents of the cholesterol hypothesis) had an undeclared conflict of interest in the form of opinions about particular diets, or through receiving funding by certain companies. Another crucial question to address is whether, and why, leading scientists disagreed about the safety of lowering cholesterol by diet or drugs.

In sum, many issues are yet to be addressed. We lacked space to discuss all the loci and reasons identified in section 2; we mainly focused on different problem definitions, on evidence of mechanisms, and on aggregating evidence from different sources. Although this contribution is far from exhausting the numerous and complex questions about scientific disagreement and, for that matter, about disagreement in the cholesterol wars, we hope to have opened up a promising line of research. Our contribution, we hope, will help in the discussion of controversies of the (distant or recent) past, and also of contemporary controversies. From the Zika outbreak to the anti-vaccine movement, from the economic crisis to climate change, there is urgent need to understand why consensus is not always secured and how this impacts policy decisions.

ACKNOWLEDGEMENTS

Earlier versions of this paper have been presented at meetings of the EBM+ consortium (ebmplus.org). We are extremely grateful to our colleagues, and particularly to Donald Gillies and Mike Kelly, for insightful and stimulating comments. We also wish to thank the editors for the opportunity to develop these ideas and an anonymous referee for urging us to clarify the distinction between ‘inconsistency’ and ‘disagreement’. Any remaining errors or inaccuracies remain, of course, ours. Research for this paper has been partly funded by the Arts and Humanities Research Council via grant AH/M005917/1 and by The Leverhulme Trust, grant id RPG-2014-181.

REFERENCES

- Anitschkow, N. (1913) Über die Veränderungen der Kanichnenaorta bei experimenteller Cholesterinsteatose. *Beiträge zur pathologischen Anatomie und zur allgemeinen Pathologie*, 56:379-404.
- Aschof, L. (1933) Introduction. In Cowdry, E. W., editor, *Arteriosclerosis*, pages 1-19. The MacMillan Company, New York.
- Bechtel, W., Richardson, R. (2010) *Discovering Complexity*. MIT Press, Cambridge, MA.
- Boon, M. (2015) The scientific use of technological instruments. In Hansson, S. O., editor, *The Role of Technology in Science: Philosophical Perspectives*, chapter 4. Springer Netherlands.
- Broadbent, A. (2011) Inferring causation in epidemiology: mechanisms, black boxes, and contrasts. In Illari, P. M., Russo, F., and Williamson, J., editors, *Causality in the sciences*, pages 45-69. Oxford University Press, Oxford.
- Brown, B. (1990) How to be realist about inconsistency in science. *Studies in History and Philosophy of Science*, 21(2):281-294.
- Brown, B. (2002) Approximate truth. In Meheus, J., editor, *Inconsistency in science*, pages 81-103. Kluwer Academic Publisher.
- Bruger, M. and Oppenheim, E. (1951) Experimental and Human Atherosclerosis: Possible Relationship and Present Status. *Bulletin of the New York Academy of Medicine*, 27(9):539.

- Chisolm, G. M. and Steinberg, D. (2000) The oxidative modification hypothesis of atherogenesis: an overview. *Free Radical Biology and Medicine*, 28(12):1815-1826.
- Clarke, B., Gillies, D., Illari, P., Russo, F., and Williamson, J. (2014) Mechanisms and the evidence hierarchy. *Topoi*, 33(2):339-360.
- Collaborators, C. T. T. C. (2005) Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90056 participants in 14 randomised trials of statins. *The Lancet*, 366(9493):1267-1278.
- Connor, W. Connor, S. (2002) Dietary cholesterol and coronary heart disease. *Current Atherosclerosis Reports*, 4(6):425-432.
- Davies, G. H. (1969) *The Earth in decay: A history of British geomorphology*, 1578-1878. Macdonald & Co.
- Duff, G. McMillan, G. C. (1951) Pathology of atherosclerosis. *The American Journal of Medicine*, 11(1):92-108.
- Eckel, R. H., Jakicic, J. M., Ard, J. D., Hubbard, V. S., de Jesus, J. M., Lee, I.-M., Lichtenstein, A. H., Loria, C. M., Millen, B. E., Miller, N. H., Nonas, C. A., Sacks, F. M., Smith, S. C., Svetkey, L. P., Wadden, T. W., and Yanovski, S. Z. (2013) *AHA/ACC guideline on lifestyle management to reduce cardiovascular risk*. *Circulation*.
- Endo, A. (2010) A historical perspective on the discovery of statins. *Proceedings of the Japan Academy, Series B*, 86(5):484-493.
- Gillies, D. (2005) Hempelian and Kuhnian approaches in the philosophy of medicine: the Semmelweis case. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36:159-181.
- Gillies, D. (2011) The Russo-Williamson Thesis and the question of whether smoking causes heart disease. In Illari, P. M., Russo, F., and Williamson, J., editors, *Causality in the Sciences*, pages 110-125. Oxford University Press, Oxford.
- Gofman, J. W., Lindgren, F. T., and Elliott, H. (1949) Ultracentrifugal studies of lipoproteins of human serum. *Journal of Biological Chemistry*, 179(2):973-979.
- Goldstein, J. L. and Brown, M. S. (2009) The ldl receptor. *Arteriosclerosis, thrombosis, and vascular biology*, 29(4):431-438.
- Gregersen, F., Køppe, S. (1988) Against epistemological relativism. *Studies in History and Philosophy of Science*, 19(4):447-487.
- Grundy, S. M., Bilheimer, C. D., Blackburn, H., Brown, W. V., Kwiterovich Jr, P. O., Mattson, F., Schonfeld, G., and Weidman, W. H. (1982) Rationale of the diet-heart

- statement of the american heart association - report of nutrition committee. *Nutrition Today*, 17(5):16-20.
- Hooper, L., Summerbell, C. D., Higgins, J. P. T., Thompson, R. L., Capps, N. E., Smith, G. D., Riemersma, R. A., and Ebrahim, S. (2001) Dietary fat intake and prevention of cardiovascular disease: systematic review. *British Medical Journal*, 322(7289):757-763.
- Howick, J. (2011) Exposing the vanities and a qualified defence of mechanistic evidence in clinical decision-making. *Philosophy of Science*, 78(5):926-940. *Proceedings of the Biennial PSA 2010*.
- Kannel, W. B., Dawber, T. R., Kagan, A., Revotskie, N., and Stokes, III, J. (1961) Factors of risk in the development of coronary heart disease - six-year follow-up experience: the framingham study. *Annals of Internal Medicine*, 55(1):33-50.
- Keys, A. (1966) Epidemiological studies related to coronary heart disease: characteristics of men aged 40-59 in seven countries. *Acta Medica Scandinavica*, 180:4-5.
- Kronmal, R. (1985) Commentary on the published results of the lipid research clinics coronary primary prevention trial. *Journal of American Medical Association*, 253(14):2091-2093.
- LaRosa, J., Hunninghake, D., Bush, D., Criqui, M., Getz, G., Gotto Jr, A., Grundy, S., Rakita, L., Robertson, R., and Weisfeldt, M. (1990) The cholesterol facts. a summary of the evidence relating dietary fats, serum cholesterol, and coronary heart disease. a joint statement by the american heart association and the national heart, lung, and blood institute. the task force on cholesterol issues, american heart association. *Circulation*, 81(5):1721.
- Lugg, A. (1978) Disagreement in Science. *Journal for General Philosophy of Science*, 9(2):276-292.
- Mabuchi, H., Haba, T., Tatami, R., Miyamoto, S., Sakai, Y., Wakasugi, T., Watanabe, A., oizumi, J., and Takeda, R. (1981) Effects of an inhibitor of 3-hydroxy 3-methylglutaryl coenzyme a reductase on serum lipoproteins and ubiquinone-10 levels in patients with familial hypercholesterolemia. *New England Journal of Medicine*, 305(9):478-482.
- Mann, G. (1985) Coronary heart disease – “doing the wrong things”. *Nutrition Today*, 20(4):12-15.
- Mccormick, J. and Skrabanek, P. (1988) Coronary heart disease is not preventable by population interventions. *The Lancet*, 332(8615):839-841.
- McMichael, J. (1979) Fats and atheroma: an inquest. *Br Med J*, 1(6157):173-175.
- Meheus, J. (2003) Inconsistencies and the dynamics of science. *Logic and logical philosophy*, 11:129-148.

- Mitchell, J. (1984) What constitutes evidence on the dietary prevention of coronary heart disease? cosy beliefs or harsh facts? *International journal of cardiology*, 5(3):287-298.
- Mitchell, S. D. (2002) Integrative pluralism. *Biology and Philosophy*, 17(1):55-70.
- Mitchell, S. D. (2003) Biological Complexity and Integrative Pluralism. *Cambridge Studies in Philosophy and Biology*. Cambridge University Press.
- Muller, C. (1939) Angina pectoris in hereditary xanthomatosis. *Archives of Internal Medicine*, 64:675-700.
- Mumpower, J. L. and Stewart, T. R. (1996) Expert judgement and expert disagreement. *Thinking & Reasoning*, 2(2-3):191-212.
- Nersessian, N. J. (2002) Inconsistency, generic modeling, and conceptual change in science. In Meheus, J., editor, *Inconsistency in science*, pages 197-211. Kluwer Academic Publisher.
- Norton, J. D. (2002) A paradox in Newtonian gravitation theory ii. In Meheus, J., editor, *Inconsistency in science*, pages 185-195. Kluwer Academic Publisher.
- Ophuls, W. (1933) The Pathogenesis of Arteriosclerosis. In Cowdry, E. W., editor, *Arteriosclerosis*, pages 249-270. The MacMillan Company, New York.
- Page, I. H., Allen, E. V., Chamberlain, F. L., Keys, A., Stamler, J., Stare, F. J., et al. (1961) Dietary fat and its relation to heart attacks and strokes. *Circulation*, 23(1):133-136.
- Peters, J. P. and Van Slyke, D. D. (1946) *Quantitative Clinical Chemistry*. Williams & Wilkins, Baltimore.
- Rifkind, B. (1984) The lipid research clinics coronary primary prevention trial results. 2. the relationship of reduction in incidence of coronary heart-disease to cholesterol lowering. *Jama-Journal of the American Medical Association*, 251(3):365-374.
- Robertson, T. L., Kato, H., Rhoads, G. G., Kagan, A., Marmot, M., Syme, S. L., Gordon, T., Worth, R. M., Belsky, J. L., Dock, D. S., Miyanishi, M., and M.D., S. K. (1977) Epidemiologic studies of coronary heart disease and stroke in japanese men living in japan, hawaii and california. *The American Journal of Cardiology*, 39(2):239-243.
- Ross, R. Glomset, J. (1976a) The pathogenesis of atherosclerosis (first of two parts). *The New England journal of medicine*, 295(7):369-377.
- Ross, R. and Glomset, J. (1976b) The pathogenesis of atherosclerosis (second of two parts). *The New England journal of medicine*, 295(8):420.
- Rudwick, M. J. S. (1969) The glacial theory. *History of Science*, 8:136.

- Russo, F. Williamson, J. (2007) Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21(2):157-170.
- Scandinavian Simvastatin Survival Study Group (1994) Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4s). *The Lancet*, 344(8934):1383-1389.
- Smith, J. M. (1988) Inconsistency and scientific reasoning. *Studies in History and Philosophy of Science*, 19(4):429-445.
- Steinberg, D. (2007) *The cholesterol wars: the skeptics vs. the preponderance of evidence*. Academic Press, San Diego, California, 1st edition.
- Stewart, T. R. (1991) Scientists' uncertainty and disagreement about global climate change: A psychological perspective. *International Journal of Psychology*, 26(5):565-573.
- Thagard, P. (1998) Explaining disease: correlations, causes, and mechanisms. *Minds and Machines*, 8:61-78.
- The Coronary Drug Project (1975) Clofibrate and niacin in coronary heart disease. *JAMA*, 231(4):360-381.
- US Department of Health and Human Services and others (1984) Lowering blood cholesterol to prevent heart disease. In *Consensus Development Conference Statement, volume 5*.
- Weiss, S. and Minot, G. (1933) Nutrition in relation to arteriosclerosis. In Cowdry, E. W., editor, *Arteriosclerosis*, pages 233-248. Macmillan company.
- Yamamoto, A., Sudo, H., and Endo, A. (1980) Therapeutic effects of ml-236b in primary hypercholesterolemia. *Atherosclerosis*, 35(3):259-266.