Medical Hypotheses (2006) **67**, 980–988

ELSEVIER

medical hypotheses

http://intl.elsevierhealth.com/journals/mehy

Popperian epidemiology and the logic of bi-conditional *modus tollens* arguments for refutational analysis of randomised controlled trials

Elard Koch [a,*], Alvaro Otarola [a], Tomas Romero [b], Aida Kirschbaum [a], Esteban Ortuzar [a]

[a] *Division of Epidemiology, School of Public Health, Faculty of Medicine, University of Chile, 939 Independencia, Santiago 70012, Chile*
[b] *Cardiology Services and Cardiac Catheterization Laboratory, Sharp CVMC San Diego, CA, United States*

Received 7 March 2006; accepted 13 March 2006

Summary Popperian epidemiology is a biomedical science tool based on the hypothesis-deductive method and the falsifiability of scientific hypotheses. This article explores the applicability of the refutationist logic tools in the analysis of a randomised controlled trial (RCT), the randomised Aldactone evaluation study (RALES). This was carried out by using bi-conditional *modus-tollens* arguments of the type (i) *P-then-*Q_n and (ii) Q_n*-If-*X_P, X_P being a set of potential falsifiers of Q_n as part of the explicit falsity-content of *P*. In this model, *P* is the main hypothesis and Q_n one or more logical predictions to be tested. The X_P argument represents inclusion criteria, exclusion criteria and conditional criteria of the RCT so every *P-then-*X_P argument should be fulfilled in canonical form to corroborate *P-then-*Q_n. Thus, falsifiability of a RCT would be determined by the empirical content of the conditional argument Q_n*-If-*X_P and its external validity would be determined by the empirical content of X_P. In this way it would be possible to mathematically assess the external validity of a RCT if the observational predicates of the X_P argument in a given population are known. According to this popperian model, applicability of the RCT results to clinical practice implies transferring of all its empirical content, in other words, the totality of its truth and falsity contents. Thus, to ignore the explicit falsity-content of a RCT such as RALES may jeopardise its potential benefits in clinical practice as suggested by recent studies.

The use of logic deductive tools in epidemiology, derived from Karl Popper's philosophy of science [1,2] in the analysis of methods [3], designs [4], statistics [5] and causal inferences [6,7] are today part of what could be called refutationist epidemiology, and probably better known as *popperian epidemiology* [8]. The refutationist approach suggests a rigorous application of the hypothesis-deductive method and a strict adherence to the principle of falsifiability as a rule for logic reasoning [6,9–11].

* Corresponding author. Tel.: +56 2 978 6433; fax: +56 2 737 1030.
E-mail address: elard@123mail.cl (E. Koch).

0306-9877/$ - see front matter
doi:10.1016/j.mehy.2006.03.033

This method has been developed in the last three decades and frequently has collided with the paradigm of the classic epidemiology based on the pragmatic justification of the induction principle and the hypothesis confirmation [8–16]. An intense debate rich in criticism [11–20], suggestions of semantic changes [10,15,21] and methodology adaptations [9,22] along with applications in observational epidemiological studies has ensued [3,23]. Recently, Hyams [24] illustrates how non-falsifiable hypotheses are insufficient to advance in medical knowledge, even when there is an abundance of inductively supported empirical data.

This article explores the applicability of the hypothesis-deductive method and the falsifiability principle in one of the main methodological designs in clinical research, the randomised controlled trial (RCT). Therefore, this study will be based on the application of the *modus tollens* syllogism to a model of a RCT using bi-conditional logical arguments [25].

Hypothesis-deductive method

Given hypothesis H along with a set of baseline conditions we deduce logical consequences $e_1, e_2, \ldots, e_n$. These consequences have to be empirically contrasted. Contrasting is the possibility of refuting H if the empirical data do not coincide with the predictions $e_1, e_2, \ldots, e_n$ derived by a logical deductive process using H as the starting point. Using a logical deductive method, observation or experimentation is subordinated to theory. In contrast, in the inductive method theory is subordinated to observation [26].

Medical scientists confront a set of initial conditions and technological problems for experimental designs in the hypothesis-deductive method, whose logical consistency is fundamental to determine the internal validity of the results reached [6]. The way of sample selection, inclusion and exclusion criteria, avoidance of biases, randomisation, data collection, follow-up strategy along with a set of auxiliary assumptions will be articulated deductively with the leading hypothesis to carry on the epidemiological study. This process assumes a general theory of the observation and experimentation [27] based on the refutationist perspective necessary for choosing among the innumerable observation objects (Fig. 1A). To conduct an experiment it is necessary a logical arrangement limited to time and space of techniques and observation procedures that the investigator considers necessary to corroborate or refute his hypothesis [28].

The inductive fallacy

In the evaluation of epidemiological hypotheses the ''introjection'', also called ''retroduction'' [10,15] of empirical data is usually confused with a kind of inductive inference (Fig. 1B and C). By definition, the induction is produced once experimental observation has been carried out and whose logical argument is a fallacy of the *modus ponens* syllogism, by means of which an antecedent statement is verified from a consequent statement, i.e.,

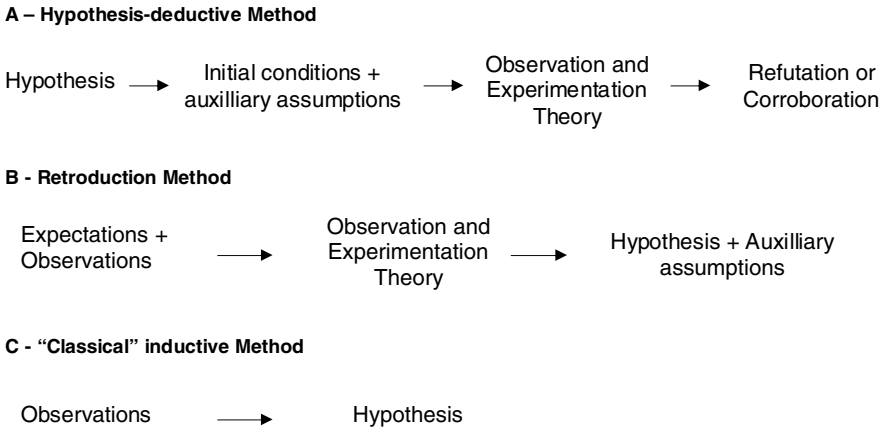

Figure 1 Deductive and inductive methodological approaches for scientific research. A and B are deductive methods, because always exist a logical ordering of observations, expectations and hypotheses, that are formulated in a theory of observation and experimentation. While in A exists a well-defined hypothesis that can be refuted, in B exists expectation that along with observation would allow to better delimit a problem or hypothesis associated to an indeterminate number of auxiliary assumptions. C represents the classic concept of induction which goes from a finite set of observation to a general hypothesis that explains them.

from the truth of particular observation e_n is concluded the truth of the general hypothesis H, which is called *fallacy of affirming the consequent* [29]. Its logical form can be formulated as:

Premise 1: H-then-e_n
Premise 2: e_n
Conclusion: H

As Greenland points out [15], this fallacy is a common practice in epidemiology. The *modus ponens* as all syllogisms is deductive and it only implies the corroboration of a hypothesis whose logical predictions have provisionally been fulfilled after an experiment that has severely tried to refute them [26]. Its correct logical form is:

Premise 1: H-then-e_n
Premise 2: H
Conclusion: e_n

From the refutationist perspective, every epidemiologic observation rationally carried out is always preceded by a theory of observation and every scientific method of experimentation is always deductive on its way to contrast subordinating the observed objects to the theoretical design of the experiment.

Falsifiability principle

According to the falsifiability principle, a theory is scientific if it can be falsified by means of experience or by its internal inconsistency [1]. A scientific statement is falsifiable when it may be logically connected to one or more observational predicates that if are not corroborated by the experiment demonstrate that the precedent statement, or hypothesis, is false. Refutation is the corroboration that an enunciation is false for not having resisted the falsification test. This procedure does not require a process of induction which is logically impossible, but a simple logical deduction. This deduction is based on the syllogism called *modus tollens* [29], illustrated as follows:

Premise 1: *P*-then-*Q*
Premise 2: not *Q*
Conclusion: not *P*

If a series of consequences derive from one hypothesis and at the same time we are able to identify a series of contradictory possibilities with those consequences, we have at our reach a series of potential falsifiers of the theory [1,26]. Examples of falsifiers are what we usually call inclusion, exclusion and conditional criteria in the design of a clinical trial. If the number of possible falsifiers of one hypothesis is greater than the one corresponding to a competing hypothesis, the first one will have more opportunities to be refuted by the experiment. In other words, it is ''falsifiable to a larger degree''. This means that the first hypothesis communicates more about the universe of the experiment than the second due to the fact that excludes a greater number of basic statements. Thus, a scientific theory has a greater degree of corroboration when it has resisted more criticisms and has been subject to more severe contrasts and not when it has been more verified [30].

Logic of bi-conditional *modus tollens* arguments

The biological phenomena are complex and are subject to multiple causal relations. Experimental hypothesis often include auxiliary assumptions and therefore the theories seldom can be absolutely refuted or corroborated [31]. On the other hand, a naïve application of *modus tollens* turns a refuted theory into unfalsifiable leading to a logical inconsistency if the theory is correctly accomplished in the future [32]. Therefore, hypotheses can be falsified only along with an unspecified number of auxiliary assumptions.

According to the falsificationist postulates, a theory contains the totality of its logical consequences [25]. Its true consequences constitute its truth-content and its false consequences, if present, its falsity-content [33]. An explicit truth-content has predictions emanating from a hypothesis that needs to be corroborated by the experiment, i.e., the beneficial effects of a drug on a certain disease. An explicit falsity-content has one or more potential falsifiers that may contradict the possible truth of a hypothesis and are defined at the onset of the experiment, i.e., the inclusion and exclusion criteria for a RCT evaluating the effectiveness of a drug. Necessarily a theory will have an unexplicited truth- and falsity-content that will be unknown and subject to investigation. Moreover, frequently in the experimental design auxiliary assumptions are defined because act as effect modifiers and potential falsifiers of one or more predictions originated from the hypothesis [22,23].

To logically interconnect the auxiliary with the main assumptions, the *modus tollens* arguments may be utilised by means of the bi-conditional propositions [25,33] such as (i) *P-then-*Q_n and (ii) Q_n-*if-*X_P, X_P being a set of Q_n's potential falsifiers

as part of the explicit falsity-content of P and Q_n being, on the other hand, logical predictions of the explicit truth-content of P. Thus, P-*then*-Q_n may be deductively refuted by the failure of Q_n-*if*-X_P and corroborated by the fulfillment of Q_n-*if*-X_P. This modality may be particularly appropriate for the analysis of double-blind RCT and others epidemiological designs.

The randomised controlled trial

According to the prevailing epidemiological theory of observation and experimentation [27,34,35], the different research designs may be organised in a hierarchical fashion in reference to their ability to establish a cause-effect relation (Fig. 2). At the present, a double-blind RCT is the closest epidemiological research model to the ideal experimentation since there is control and manipulation of the dependent and independent variables. Maclure [4,23] comments that from the refutationist view point, the randomisation and blinding of a RCT minimise the selection and information bias and the non-chance confusion with a non-experimental or observational study in which is necessary to refute many competing hypothesis before a causal relation can be inferred. The blinding intervention minimises the potential subjective interferences of the researcher and the patient when a treatment is compared to a placebo. The strict adherence to the inclusion and exclusion criteria makes possible the control of possible biases and variables that may modify the results. In this fashion it is also possible to test one or more main hypotheses at the same time. From the refutationist point of view, the RCT is a rigorous falsification test utilising the hypothesis-deductive method to control a greater number of potential falsifiers related to one or more explicitly formulated main hypotheses.

Despite that RCTs are thought to have the highest value in epidemiological and clinical research, they certainly have some limitations [36—41]. By using strict inclusion, exclusion and conditional criteria, numerous subjects are often excluded. The more complex a RCT, the more rigorous selection, limited inclusion of subjects and more controlled intervention is carried out. Thus, both the population and the environment of the experimental intervention are often different from those encountered in clinical practice [39—42]. Indeed these restricted conditions of the experiment increase its internal efficacy but negatively affect its external validity [41]. Moreover, the RCT groups are created by randomisation of a given population but not by random sampling within the groups of the population under study. This may lead to errors in the extrapolation of statistical inferences to the general population.

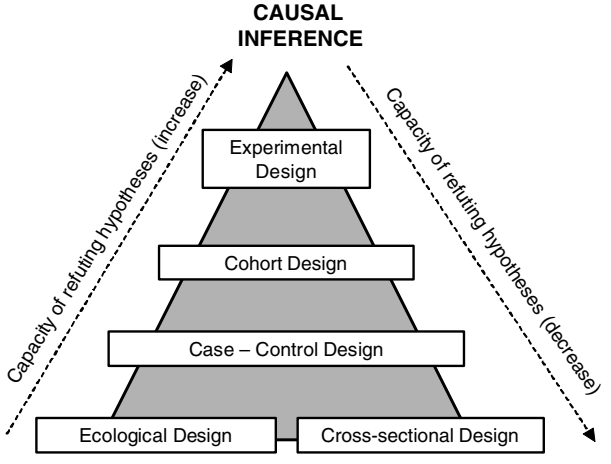

Figure 2 A pyramidal scheme for a directional classification of the study designs in epidemiological research. From a hierarchic model, the strength to confirm cause-effect association increases from the observational level to the experimental level, a traditional concept of the inductive epidemiology. From the refutationist perspective, the capacity of refuting causal hypotheses is greater for experimental design and smaller for cross-sectional uncontrolled studies.

A multinational clinical trial model

Today special attention is granted to the results provided by large multinational RCTs in contrast to smaller studies. Their main characteristic is that the subjects composing the samples are recruited from different parts of the world [42]. The weight of evidence emanating from these larger RCTs is often considered decisive in the assessment of a new therapeutic intervention since they provide a greater ''inductive support'' to generalise their results to the population beyond the limits of the study. This may unduly influence the introduction of a new prescription drug in clinical practice [43].

A good example of a large multinational RCT is the randomised Aldactone evaluation study (RALES). This study was published by Pitt et al. [44] and involved 195 centres in 16 countries. A number of 1663 patients with severe congestive heart failure (CHF) constituted the population studied, 822 randomly assigned to receive 25 mg of spironolactone and 841 to placebo. Four countries enrolled more than 100 patients ($n = 1138$) and 12 countries enrolled less than 100 patients ($n = 525$). All the patients had significant left ventricular dysfunction as defined by two-dimensional echocardiography. The

study concluded that spironolactone, in addition to the usual treatment that included diuretics and angiotensin-converting enzyme (ACE) inhibitors, significantly decreased, in a period of two years, all-cause mortality by 30%, sudden death by 29% and hospitalisation related to CHF by 35% when compared to placebo. The incidence of hyperkalaemia was minimal both in the treated and placebo groups (2%). The authors concluded that the blockade of aldosterone receptors along with the standard therapy in patients with severe CHF secondary to left ventricular systolic dysfunction reduced significantly both the morbidity and mortality with minimal negative side-effects. RALES was stopped earlier than the anticipated length of the study because an interim analysis determined significant beneficial effects, leading to a rapid diffusion of its results along with an explosive increase in the use of spironolactone after its publication [45].

Logical arguments

The main explicit hypotheses of the authors of RALES were (a) the use of low doses of Spironolactone along with the standard therapy would significantly reduce both mortality and morbidity in patients with severe CHF and (b) this would occur without major negative side effects, mainly hyperkalaemia. On these premises a general proposition can be defined as: P, ''Spironolactone is better than placebo in patients with severe CHF''. Then, three consequent propositions derived from P can follow: Q_1, ''Spironolactone reduces mortality in patients with severe CHF''; Q_2, ''Spironolactone reduces CHF related morbidity''; Q_3, ''Spironolactone produces minimal complications because of hyperkalaemia''. Using the *modus tollens* syllogism, the hypothesis can be enunciated as P-*then*-Q_n, where all the predicates components of Q_n need to come true to corroborate the RALES hypothesis; they are potential falsifiers that are part of the explicit truth-content of P.

Traditionally the refutationist epidemiology have utilised as logic models single *modus tollens* arguments for each auxiliary assumption competing with the main hypothesis such as A-then-Q, B-then-Q, etc. However, it is possible to use bi-conditional arguments to connect a set of auxiliary assumptions with the main hypothesis and its predictions. As mentioned above, RALES included a number of inclusion, exclusion and conditional criteria that can be described as X_P (Table 1). These auxiliary assumptions if not fulfilled can refute the logical prediction P-*then*-Q_n as part of the explicit falsity-content of P. In the design of a RCT these observational predicates are part of the antecedent statement prior to the formulation of the logical prediction Q_n. In other words, before the prediction P-*then*-Q_n comes true, P-*then*-X_1, P-*then*-$X_2, \ldots, P$-*then*-X_n have to be accomplished all of which can be expressed as a *modus tollens* single argument, P-*then*-X_P. However, this type of argument, P-*then*-X_P and X_P-*then*-Q_n constitutes a fallacy by denying the antecedent in X_P. To resolve this problem the premise can be formulated as a bi-conditional argument:

(i) P-*then*-Q_n and,
(ii) Q_n-*if*-X_P

The argument (ii) is shaped by transferring the explicit falsity-content of the precedent statement P to each one of the consequent statements Q_n. The hypothesis that the placebo is challenging the main hypothesis can be expressed as P'-*then*-Q'_n. Finally, the expected result is that Spironolactone is better than placebo, or $P > P'$ so $Q_n > Q'_n$ should come true. According to this model, an inclusion or exclusion criteria is a potential falsifier of the hypothesis under corroboration when:

(1) It behaves as a competing hypothesis such as proposition C, in which case the prediction C-*then*-Q_n competes with P-*then*-Q_n so it cannot be isolated the effect of P (selection bias, i.e., X_1 to X_3, Table 1).
(2) It can simultaneously falsify P and Q (a confounding factor , i.e., X_5 and X_{14}, Table 1).
(3) It may falsify Q_n independently of P (i.e., X_6 and X_{13}, Table 1).
(4) It may directly falsify P (information bias, i.e., non-blinded study).

Theoretical refutation of the RCT model

It is possible to demonstrate using canonic logic arguments the falsifiability of RALES. For example, we can assume that ''minimal complications by hyperkalaemia'' is false, or *no*-Q_3, and that the exclusion of ''diabetes'' or predicate X_4 (Table 1) was not accomplished. This situation can be expressed as a complete canonical logical argumentation: (i) if P-*then*-Q_3 and (ii) Q_3-*if*-X_P, X_P being (iii) the set $X_1, X_2, X_3, \ldots, X_n$ and (iv) *no*-X_4, (v) *no*-Q_3 then (vi) *no*-P. By deduction, *no*-X_P can be any inclusion, exclusion or conditional criteria of RALES that was not fulfilled in any other population. Therefore, if any of these auxiliary assumptions contained in X_P is false the hypothesis P would be to some extent refuted [46].

Table 1 Inclusion criteria, exclusion criteria and conditional criteria as part of the explicit falsity-content (X_P) of RALES experimental design. (Source: Ref. [46], Copyright 2005 BMJ Publishing Group. All rights reserved.)

X_P arguments	Observational predicates	Criteria
X_1	Patients having suffered chronic heart failure IV degree of the New York Heart Association during the 6 previous months	Inclusion
X_2	Functional class III or IV at the moment of inclusion, diagnosed as cardiac insufficiency at least 6 weeks earlier	Inclusion
X_3	Patients with an ejection fraction of the left ventricle below 35% during the last 6 months	Inclusion
X_4	Diabetes	Exclusion
X_5	Renal insufficiency or creatinine above 2.5 mg/dl	Exclusion
X_6	Congenital cardiopathy	Exclusion
X_7	Unstable angina	Exclusion
X_8	Patients with potassium > 5 mm/l	Exclusion
X_9	Hyperkalaemia history	Exclusion
X_{10}	Operable valvular disease	Exclusion
X_{11}	Primary hepatic failure	Exclusion
X_{12}	Cardiac transplant indication	Exclusion
X_{13}	Active cancer or any life-threatening disease	Conditional
X_{14}	Close laboratory monitoring during treatment	Conditional (research process)
X_{15}	Potassium supplements were not recommended, except when potassium levels in serum were below 3.5 mmol/l	Conditional (research process)
X_{16}	Patients had to be in treatment with a loop diuretic and an angiotensin-converting enzyme inhibitor if it were tolerated	Conditional (research process)
X_{17}	Treatment with digitalis and vasodilators was allowed but not with potassium sparing diuretics	Conditional (research process)

According to the falsificationist theory, when the results of a RCT like RALES are extrapolated to the clinical practice all its empirical content is transferred, in other words, all its truth- and falsity-content [46]. The non-fulfillment of the explicit falsity-content of a RCT may lead to a cancellation of its potential clinical benefits. For example, Bozkurt et al. [47] studied the outcomes

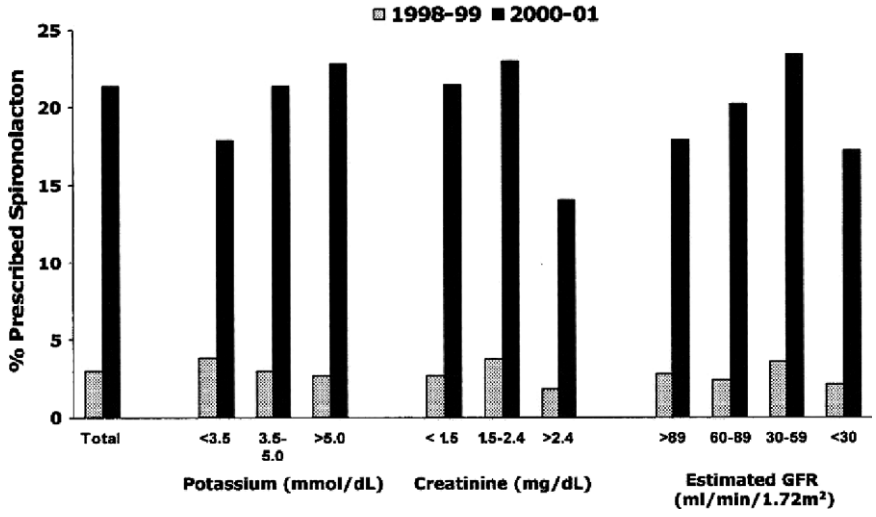

Figure 3 Changes in proportions of patients receiving prescription of spironolactone at hospital discharged from 1998 to 1999 (before RALES) and from 2000 to 2001 (after RALES) in all patients and patients stratified by serum potassium, creatinine, and estimated glomerullar filtration rates (GFR). (Source: Ref. [48].)

of 377 patients started on Spironolactone for CHF after the publication of RALES. They found that, in this ''real world'' study, few patients fulfilled the inclusion and exclusion criteria of RALES, nor its outcomes. For example, 40% of the patients were diabetics, 24% developed hyperkalaemia, severe in 12% (>6 mmol/l). A main conclusion in this study was that results obtained from multi-centre large RCT may lead to a false impression of safety by ''inductive support'' with the consequent injudicious use therapies and increase in adverse outcomes. Recently, Masoudi et al. [48] observed that the prescription of spironolactone increased from 3% to 30% in older adults with CHF in U.S. after RALES publication. Moreover, 30.9% was provided to patients not meeting enrollment criteria of the study. In the population of patients discharged from 2000 to 2001, spironolactone was prescribed to 22.8% of patients with a serum potassium value $\geqslant$5.0 mmol/l, to 14.1% of patients with a serum creatinine value $\geqslant$2.5 mg/dl, and to 17.3% of patients with severe renal dysfunction (Fig. 3).

Discussion

On the philosophical debate the refutationist epidemiology has been shown in opposition to the scientific paradigm of inductive epidemiology [8–14]. The controversy has been centred on the problem of induction which according to Greenland [15] would be of semantic order based on the radical anti-inductivism that Karl Popper maintained all his life [49]. However, since Hume [50] formulated the problem in the XVIII century, it does not exist a logically valid form that justifies the principle of induction. It is only possible to accept its pragmatic justification [16,51] as part of practical decision making on the formulation of rational predictions and transference of results to the population, a question that keeps controversy between epidemiologist [10,11,28,52]. According to the current refutationist approach [11,19,27,28,46], every hypothesis has only one value of provisional truth related to its degree of corroboration and testability that would allow to act over the ''real empirical world''. It is not possible to correct errors using an inductive method, because *modus ponens* syllogism only allows to deductively corroborate a prediction assuming the truth of an initial hypothesis. Finally, the correction of an error always implies the falsification of a hypothesis associated to an indeterminate number of auxiliary assumptions, which is deductively carried out by means of *modus tollens* syllogism. Indeed, statistical tests used in epidemiologic research are based on refutation of the null hypothesis and the probability of a significant test is, in practical terms, inseparable from the quality of hypothesis tested, especially in multivariate models [5]. Thus, epidemiology does not progress by induction or verification of hypotheses, but by correction of mistakes in previous theories through an essay and error method [30,31,49].

It has been argued in favour of the induction that probabilistic statements would be unfalsifiable [8,13,53]. However, Popper and Miller demonstrated the impossibility of an *inductive probability* [54] and on the contrary, it is a mathematical reasoning absolutely deductive [55]. On the other hand, the denial of the falsifiability principle from the uncertainty principle leads to a sophism whose logical consequence will be the denial to every necessity of experimentation in probabilistic terms. If a scientific statement whether probabilistic or categorical is unfalsifiable: Why is it necessary to experiment in epidemiology or in other sciences that base their premises on probabilistic statements? One of the fundamental characteristics of scientific knowledge, mainly in biomedical sciences, is its fallibility [56]. In epidemiology this characteristic springs to the eye when it is observed that a large part of the probabilistic results are instantaneously refuted by the common experience [22,28]. For example, if smoking is associated with a higher probability to contract lung cancer, it will be observed that not all the people who smoke contract cancer. However, by logical synthesis, the general scientific statement that contains the probabilistic statements of a causal relation like this is ''*Smoking produces lung cancer*'' and to refute this hypothesis requires studies that corroborate a higher probability of non-smokers to get lung cancer. Thus, the causal association between smoking and lung cancer is nowadays a well-corroborated hypothesis from the method that Mill [57] called *inverse deduction* [58] which is common in all empirical sciences [59].

The falsifiability principle of the scientific enunciations is the logical fundament which justifies the necessity of experimenting through hypotheses [24,28,60]. In this study we propose the application of the falsifiability principle to a multi-site RCT as RALES. The epidemiological experimental design is well adapted to the logical structure of bi-conditional *modus tollens* arguments allowing to carry out an axiomatic refutation such as (i) *P-then-*Q_n and (ii) Q_n*-If-*X_P, X_P being a set of potential falsifiers of Q_n as part of the explicit falsity-content of *P*. In this model *P* represents the hypothesis and Q_n one or more logical predictions that have to be empirically con-

trasted, while X_P represents the criteria of inclusion, exclusion and conditional criteria that have to be canonically fulfilled during the experimentation for the corroboration of *P-then-*Q_n. Then, the falsifiability of a RCT would be determined by the empirical contents of each conditional argument Q_n-*if-*X_P. However, considering an axiomatic refutation like the one carried out in the RALES model and its testability demonstrated in a canonical form: Which is the degree of falsifiability in quantitative terms? Is it possible a probabilistic approximation to the generalisability of a RCT knowing the probability of the bi-conditional arguments? Certainly, an important problem in epidemiologic research is the external validity of RCTs [41]. Recently, in a companion paper [46], we have analysed the results of a populational study carried out in Ontario, Canada [61] that has questioned RALES predictions showing at the same time the difficulties of the ''inductive support'' provided by great multinational RCTs as a criterion of external validity. It would be the probabilistic knowledge of the observational predicates contained in X_P and not the selection of patients from multiple countries what would allow to estimate the generalisability of RCT for a specific population, an application whose extents should be investigated. Indeed, this logical structure allows to evaluate a highly improbable hypothesis by increasing the number of its potential falsifiers. Thus, the refutational model described can have other applications in biomedical sciences, specially facing epidemiological problems which have not been solved by the current inductive method [24,62].

Acknowledgments

We express our gratitude to Mireya Hernández for her assistance in the manuscript preparation. E. Koch is partially supported by a doctoral fellowship from MECESUP UCH-0219, School of Public Health, Faculty of Medicine, University of Chile and by a research grant from ''Fundación Araucaria'', San Diego, California.

References

[1] Popper KR. The logic of scientific discovery. London: Hutchinson; 1997.
[2] Buck C. Popper's philosophy for epidemiologists. Int J Epidemiol 1975;4:159—68.
[3] Weed D. An epidemiological application of Popper's method. J Epidemiol Community Health 1985;39:277—85.
[4] MaClure M. Taxonomic axes of epidemiology study designs: a refutationist perspective. J Clin Epidemiol 1991;44: 1045—53.
[5] Senn SJ. Falsificationism and the clinical trials. Stat Med 1991;10:1679—92.
[6] Weed D. On the logic of causal inference. Am J Epidemiol 1986;123:965—79.
[7] Renton A. Epidemiology and causation: a realist view. J Epidemiol Community Health 1994;48:79—85.
[8] Karhausen L. The poverty of popperian epidemiology. Int J Epidemiol 1995;24:869—74.
[9] MaClure M. Popperian refutation in epidemiology. Am J Epidemiol 1985;121:343—50.
[10] Poole C. Induction does not exist in epidemiology, either. In: Rothman KJ, editor. Causal inference. Chestnut Hill, MA: E.R.I.; 1988.
[11] Banegas JR, Rodriguez F, Del Rey Calero J. Popper and the problem of induction in epidemiology. Rev Esp Salud Publica 2000;74:327—39.
[12] Jacobsen M. Against popperized epidemiology. Int J Epidemiol 1976;5:9—11.
[13] Pearce NE, Crawford-Brown DJ. Critical discussion in epidemiology: problems with the Popperian approach. J Clin Epidemiol 1989;42:177—84.
[14] Ng SK. Does epidemiology need a new philosophy? Am J Epidemiol 1991;133:1073—7.
[15] Greenland S. Induction versus Popper: substance versus semantics. Int J Epidemiol 1998;27:543—8.
[16] Salmon W. The pragmatic justification of induction. In: Swinburne R, editor. The justification of induction. Oxford: Oxford University Press; 1974.
[17] Goodman SN. Toward evidence-based medical statistics 1: The *P* value fallacy. Ann Intern Med 1999;130:995—1004.
[18] Susser M. What is a cause and how do we know one? A grammar for pragmatic epidemiology. Am J Epidemiol 1991;133:635—48.
[19] Cole SR. Atheoretical science. A response to the poverty of Popperian epidemiology. Int J Epidemiol 1996;25:899—900.
[20] Maclure M. On the logic and practice of epidemiology. Am J Epidemiol 1987;126:554.
[21] Poole C. Commentary: Positivized epidemiology and the model of sufficient and component causes. Int J Epidemiol 2001;30:707—9.
[22] Maclure M. Multivariate refutation of aetiological hypotheses in non-experimental epidemiology. Int J Epidemiol 1990;19:782—7.
[23] Maclure M. Demonstration of deductive meta-analysis: ethanol intake and risk of myocardial infarction. Epidemiol Rev 1993;15:328—51.
[24] Hyams KC. The investigation of chronic fatigue syndrome: a case-study of the limitations of inductive inferences and non-falsifiable hypotheses in medical research. Med Hypotheses 2003;60:760—6.
[25] Watkins J. Popperian ideas on progress and rationality in science. Critical Rationalist 1997;2:1—11. Available from: http://www.eeng.dcu.ie/~tkpw/tcr/volume-02/number-02/v02n02.pdf.
[26] Popper KR. Conjectures and refutations: The growth of scientific Knowledge. London: Routledge; 1992.
[27] Weed D. Theory and practice in epidemiology. Ann NY Acad Sci 2001;954:52—62.
[28] Maclure M. Karl Popper and his unending quest: an epidemiologic interpretation. Epidemiology 1995;6:331—4.
[29] Blackburn S. The Oxford dictionary of philosophy. New York: Oxford; 1994.
[30] Popper KR. Replies to my critics. In: Schilpp PA, editor. The philosophy of Karl Popper. Illinois: Open Court; 1974.

[31] Lakatos I. The methodology of scientific research programmes. London: Cambridge University Press; 1978.
[32] Freyerabend P. Against method. London: New Left; 1975.
[33] Miller DW. Critical rationalism: a restatement and defense. Chicago: Open Court; 1994.
[34] Rothman KJ. Epidemiology. An introduction. New York: Oxford University Press; 2002.
[35] MacMahon B, Trichopoulos D. Epidemiology: principles and methods. Philadelphia: Lippincott Williams and Wilkins; 2001.
[36] Benson K, Hartz AJ. A comparison of observational studies and randomized controlled trials. N Eng J Med 2000;342: 1878–86.
[37] Holmberg L, Baum M, Adami HO. On the scientific inference from clinical trials. J Eval Clin Pract 1999;5:157–62.
[38] Sackett DL, Richardson WS, Rosemberg W, Haynes RB. Evidence-based medicine: how to practice and teach EBM. Edinburgh: Churchill Livingstone; 2000.
[39] Cohen AM, Stavri PZ, Hersh WR. A categorization and analysis of the criticisms of Evidence-Based Medicine. Int J Med Inform 2004;73:35–43.
[40] Naylor CD. Grey zones of clinical practice: some limits to evidence based medicine. Lancet 1995;345:840–2.
[41] Rothwell PM. External validity of randomised controlled trials: To whom do the results of this trial apply? Lancet 2005;365:82–93.
[42] Charlton BG. Fundamental deficiencies in the megatrial methodology. Curr Control Trial Cardiovasc Med 2001;2: 2–7.
[43] Ghali WA, Cornuz J. Early uptake of research finding after fast-track publication. Lancet 2000;355:579–80.
[44] Pitt B, Zannad F, Remme WJ, et al. The effect of spironolactone on morbidity and mortality in patients with severe heart failure. Randomized Aldactone evaluation study investigators. N Eng J Med 1999;341: 709–17.
[45] Majumdar SR, McAlister FA, Soumerai S. Synergy between publication and promotion: comparing adoption of new evidence in Canada and the United States. Am J Med 2003;115:467–72.
[46] Koch E, Otarola A, Kirschbaum A. A landmark for popperian epidemiology: refutation of the randomised Aldactone evaluation study. J Epidemiol Community Health 2005;59: 1000–6.
[47] Bozkurt B, Agoston I, Knowlton A. Complications of inappropriate use of spironolactone in heart failure: when an old medicine spirals out of new guidelines. J Am Coll Cardiol 2003;41:211–4.
[48] Masoudi FA, Gross CP, Wang Y, et al. Adoption of spironolactone therapy for older patients with heart failure and left ventricular systolic dysfunction in the United States,1998-2001. Circulation 2005;112(1):39–47.
[49] Mosterin J. Entrevista con Karl Popper. Episteme 2002;22:95–118. Available from: http://www.revele.com.ve//pdf/episteme/vol22-n1/pag95.pdf.
[50] Hume D. An inquiry concerning human understanding. Oxford: Oxford University Press; 1974.
[51] Putnam H. The ''*corroboration*'' of theories. In: Schilpp PA, editor. The philosophy of Karl Popper. Illinois: Open Court; 1974.
[52] Pavesi M. Induction versus Popper: substance versus semantics. Int J Epidemiol 1999;28:360–1.
[53] Greenland S. Probability logic and probabilistic induction. Epidemiology 1998;9:322–32.
[54] Popper KR, Miller DW. A proof of the impossibility of inductive probability. Nature 1983;302:687–8.
[55] Popper KR, Miller DW. Why probabilistic support is not inductive. Phil Trans R Soc Lond A 1987;321:569–91.
[56] Popper KR. The myth of framework: In defense of science and rationality. London: Routledge; 1982.
[57] Mill JS. A system of logic: ratiocinative and inductive. In: Robson JM, editor. Collected works of John Stuart Mill. Toronto: Toronto University Press; 1973.
[58] J.S. Mill called inverse deduction the method through which a set of repeated observed trends is reduced to a general theory which contains them. It was originally formulated in the field of social sciences, but Popper extended it as a common process to all the empirical sciences and he called it *method of reduction* [59]. A current example is the epidemiologic transition which was elaborated to explain the mortality patterns around the world [63].
[59] Popper KR. The poverty of historicism. London: Routledge; 1961.
[60] Antiseri D. Philosophers, physicists and physicians in defense of the unity of the scientific method. Clin Ter 1998;149:429–33.
[61] Juurlink DN, Mamdani M, Pharm D, et al. Rates of hyperkalaemia after publication of the randomized Aldactone evaluation study. N Eng J Med 2004;351:543–51.
[62] Maclure M. Inventing the AIDS virus hypothesis: an illustration of scientific vs unscientific induction. Epidemiology 1998;9:467–73.
[63] Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. Circulation 2001;104:2746–53.