Chapter II

THE LITERATURE REVIEW

Gilles Landrivon, Pierre Duhaut, Hélène Pellet

Access to information is the essential step in the development and implementation of the research question. Has this question already been considered, or is there data in the literature leading to modifying the research question? It is therefore important to find in the literature the relevant information concerning the research question, to ensure that this information is exhaustive, then to analyze it, which will be the subject of a "critical reading" (critical appraisal).

I. DOCUMENTARY RESEARCH

The documentary research must be **exhaustive** (not leaving out any article useful to the research question) and **relevant** (avoiding as much as possible articles of no interest for the research). It is done in several stages. It's necessary :

- To know what you are looking for. To do this, the subject must be defined: evolution of a disease, study of factor(s) modifying this evolution, etc..., specify the context (type of patients concerned, age, social and geographical conditions, etc.) and set the limits over time of documentary research.
- Choose the appropriate documentary sources
- Develop the research strategy, and implement it.

A. Documentary sources

They are not mutually exclusive (several sources may in fact contain a significant number of identical journals). It is therefore important to first use the most efficient documentary source in its field, before broadening the search to other sources if necessary.

1. Internet research

General search engines (Google, Yahoo, etc.), metasearch engines (Copernic), directories (Evidence-Based Healthcare, CISmef, etc.) or selective tools (SUM-search) are used to query several sources simultaneously. We call "EBM filters" (Evidence-Based Medicine) the tools offering predefined search strategies, and using more qualitative criteria. For a prolonged work, it is generally necessary to use to one or more sometimes paying databases to obtain full text articles.

a- PubMed

The most effective solution in Life and Health Sciences consists in using PubMed (www.pubmed.org), Medline's Internet interface to which access is free. Developed by the National Library of Medicine (Bethesda, USA), there were 16.8 million references in

December 2007 (the oldest of which dates back to 1865). Five hundred thousand references on average are added each year. The majority come from the Medline database, covering the period 1966 to the present day, and PubMed also provides access to the 'OldMedline' database, covering the period 1950-1966. 5200 periodicals from 80 countries are indexed. Most of the journals referenced are in English, even when they come from non-English-speaking countries, such as some high-level Scandinavian or Italian journals.

In addition to being free, PubMed combines many advantages:

- The referenced journals are above all **peer-reviewed journals**, publishing articles reviewed by experts recognized by the journal as competent in the field: this is the 'peer-review'.
- The article search strategies offered by PubMed are effective, thanks in particular to a **rich thesaurus** and the possibilities of combining these keywords.
- PubMed provides access to **English summaries** of virtually all original referenced articles (less often for general reviews or case reports).
- PubMed provides access to **full-text articles** through two main routes. The full text can be part of the PubMed article bank, and its access is then free. It can also be put online by the journal, and PubMed associates with each reference the link to the journal and the full text of the article. Two scenarios can then arise: the article is put online for free by the journal (this is the case for some major international journals, for articles published for more than 6 months, and for articles with significant repercussions in the field in question, or the article may only be accessible by paid subscription to the journal, or by online purchase of the article in question.

When articles are available online, they are often available in two formats. The **pdf format** (Adobe), (printable format for easy reading) and the **html format**, with separate text, graphics, and illustrations, which makes it easy to save graphs and illustrations and integrate them into a PowerPoint-type presentation. Some journals even provide ready-made slides with text and references for their illustrations, which can be integrated as is into a presentation.

- PubMed often provides **bibliographic references** of articles found, with the hypertext link to the corresponding article and, where applicable, to its full text if it is part of the PubMed article bank or if it is put online by the original journal. Exploring these references allows you to find interesting articles not listed during a first search.
- More and more articles are put online before they are printed and published in print: they are then referenced with the mention [**Epub ahead of print**], and can sometimes be accessible 6 months before their print publication.
- Finally, more and more institutions (universities, research structures, hospitals) currently have **collective electronic subscriptions** made available to their members (students, professionals). The publications are then accessible via specialized sites often grouping together all the publications of a publishing house, and these sites are provided with hypertext links by PubMed. Access to the original article is under these conditions fast and free for the user, if he accesses the site via a computer belonging to the network of the subscriber institution.

b. Two "general" tools: Current Contents, Cochrane Library

The Current Contents and Cochrane Library are available on the Internet, and also on CD-Rom, but with the disadvantage of annual updates.

The list of references can be printed, and the articles that one excludes can be crossed off this list.

The **Current Contents** produced by the *Institute for Scientific Information* (www.isinet.com) provide two publications (out of nine) particularly suited to clinical research, published weekly:

- CC "Life Sciences" provide the summary of 1350 periodicals,
- CC "Clinical Medicine" the summary of 1120 periodicals, some of which are in the 2 publications.

Some sources provide the references without qualitative selection. This is the case of Embase Medline and Pascal.

EBM sources make a qualitative selection. This is the case of the Cochrane Library:

- DARE (Database of Abstracts of Reviews of Effectiveness)
- CAT (Critically Appraised Topics).

c- Predefined search strategies

Clinical Queries usable on Pubmed allows you to carry out a specific search. The disadvantage, however, is the risk of missing interesting references.

When the research concerns a particular theme, it is a good idea to contact specific databases, for example:

- for **public health**: the BDSP (French Public Health Data Bank, (www.bdsp.tm.fr), free access, is managed by the National School of Public Health. It contains approximately 15,000 indexed documents with a Thesaurus.
- for **medical economics**:
  - ♣ EED (National Health Service Economic Evaluation Database)
  - ♣ Ecosanté for digital data in France
  - ♣ The CRD (Centre for Reviews and Dissemination of York) (http://www.crd.york.ac.uk/crdweb/) presents bases for economic evaluation: NHS Economic Evaluation Database and summaries of cost/cost analyses, benefit and cost/effectiveness of medical practices: EED database (NHS Economic Evaluation Database) which can be consulted free of charge.
  - ♣ The CODECS database (Knowledge and Decisions in Health Economics, (http://infodoc.inserm.fr/codecs/codecs.nsf), managed by the College of Health Economists and INSERM, reports the studies of health economics made in France.

2. Searching for articles in a library

Usually, university libraries allow researchers free access to periodicals. The researcher can therefore directly consult the periodicals, and work on the spot with his list of references, quickly eliminating the articles of no interest for his study, noting on a sheet the few elements to remember, or photocopying the articles on which he will have to work more extensively.

3. Unpublished documentary research

This aspect must be taken into account. If knowledge in medicine or in other sectors of scientific activity is most often published, the non-publication of data is not the exception. In the industrial sector, the results of clinical studies, when they are negative, do not arouse enthusiasm on the part of promoters. This is true for negative therapeutic trials for a molecule tested. However, the data produced does exist. Some subjects are politically sensitive and, for example, studies and epidemiological data relating to contamination, modes of transmission

and methods of preventing infectious diseases are not all available in real time...or those that are published are sometimes to be analyzed taking publication bias into account!

These “unpublished” data constitute what is called the “grey literature". They require institutional and professional knowledge to know how to look for data in the “right place”.

From a qualitative point of view, they can be significant, in particular in the field of new technologies (“technology assessment” studies). The development of new screening, diagnostic and treatment technologies is the subject of investigations, in the form of industrial reports or clinical pre-reports which are not always available in the initial phases of development of these technologies.

B. The documentary research strategy

- **It is first necessary to determine the keywords of the search** by using terms (or descriptors) belonging to the thesaurus in English MeSH (available on the MeSH of Browser of Pubmed, http://www.nlm.nih.gov/ mesh/).

You can use the "HONselect" interface of the Swiss foundation Health On the Net (HON) for this, which allows a quick connection and provides the French translation of the MeSH terms.

The request must be specified using *Boolean operators*: AND, OR, NOT.

Using Pubmed requires practice, and can be confusing at first.

The Paris URFIST provides a good user manual (http://www.hon. ch/HONselect/index_f.html), including an animation and corrected exercises.

- **The selection of articles** is made according to the titles, then according to the abstracts, which can be read on the computer. A good quality summary generally corresponds to a good quality article, and allows the reader to judge whether this article provides him with useful information for his study. The rapid reading of the article takes into consideration the objective of the article, the quality criteria, the reported results. The article is printed if it is considered relevant, the in-depth reading being done on paper.

C. The use of the documents collected

Rejected articles are removed from the reference lists – after consulting the references they contain. The other articles will be classified in two groups: those which present a specific interest, and those which include many interesting data.

Within each of these two groups, we will choose a method of classification, for example by name of the first author (and year, if the same name is found several times, which is frequent).

For the "specific interest" group, note the nature of this point on the list, next to the reference.

- **Control of the completeness of the search**

It is done through a dual approach: retrospective and prospective.

Retrospective control: it consists of consulting the bibliographical references of all the articles collected (including those that have been considered to be rejected) to ensure that an

interesting "historical" reference has not been missed. This work, which is essential, is relatively easy and quick. After which, one can definitively reject the articles without interest for the research question.

The prospective approach consists of identifying new publications appearing during the course of the study. This is done using alert lists. There are many possibilities: most periodicals offer free receipt by e-mail of the table of contents the same day of publication. Access to the articles can be free, as for some articles of the British Medical Journal (www.bmj.com), it can be done by a "pay-per-view" system (paid consultation), for example for the New England Journal of Medicine (www.nejm.com) or many other journals, or be restricted to subscribers.

It is easy for many newspapers to obtain the titles and abstracts of articles corresponding to selected keywords, especially in the case of periodicals indexed in databases. This service is free with MyNCBI (http://www.ncbi.nlm.nih.gov/entrez/login.fcgi?call=so.signon.login). There are also free, sponsored services, some of which also provide article conclusion (www.mdlinx.com).

- Determination of the relevance of the selected documents

The work of compiling data is necessary but not sufficient, since it must be subjected to a “critical” reading in the sense of “critical appraisal” of the Anglo-Saxons.

This is the trickiest part of the documentation. It must answer two questions:
- Is the article, and the information it contains, relevant to my research question?
- Is the article reliable? The answer to this question is provided by the critical analysis of the medical literature. In other words, are the results provided of good quality so that I can use them to construct my research question and benefit my research protocol?

II. ASSESSMENT OF THE RELEVANCE OF THE LITERATURE

This is the task of the future investigator, who wishes to clarify his clinical research question or find the justification for his draft protocol. This is also what the clinician does, seeking, through regular reading of the medical literature, decision support for his daily practice.

Reading literature requires selection and evaluation. For this, the concept of "critical reading" has been developed. The principle is to judge the value of the publications, whether it is the quality of the research undertaken or the relevance of the results published:

- What is the credibility of the publication (internal validity)?

Do the results reported by the author really correspond to reality? Can we trust the conclusions proposed by the author? To assess the validity of the study, the reader must be able to quickly identify, depending on the type of study, the different stages of the protocol that underpinned it and their components, evaluating them so as to define the level of credibility of the information provided.

- What is the applicability of the information contained in the publication (external validity)?

If the conclusions are considered valid, are they applicable to the medical practice of the reader or to the research project of the future investigator?

A. General description of the method

This reading method is based on the use of a single, standardized evaluation grid. The proposed analysis plan has the advantage of being applicable to all types of publications. It is an adaptation of the "Critical Appraisal Worksheet" of the Center for Clinical Epidemiology and Biostatistics (Pr R. F. Heller) of the University of Newcastle (New South Wales, Australia).

This grid consists of 8 lines, corresponding to 8 evaluation criteria.

Each of these 8 criteria calls for the same 3 types of questions, which correspond to the 3 columns of the grid:

1- Is it possible to find in the article the information for the criterion in question?

2- Is the way in which the criterion in question has been approached correct?

3- If the approach to the criterion in question is incorrect, does this threaten the validity of the study?

These 8 lines correspond to the main steps in designing a protocol. All publications are the result of a protocol-defined study. The analysis of the protocol makes it possible to validate it, using the grid. In responding to the eight questions, the reader has the ability to very quickly dismiss what is invalid. He can thus take an objective look at the quality of the results offered to him.

B. The eight steps

1. What is the goal?

The doctor is looking for scientific information concerning his four main concerns: therapy, prognosis, etiology and diagnosis. (These are the four categories that are offered in the clinical questions section of the NCBI PubMed online database. www.pubmed.org).

The impact of an intervention.

The medical intervention is most often therapeutic, medicinal or not, but can also be diagnostic, screening or educational. The objective is to distinguish the useful intervention from that which is useless or even dangerous. The questions to which the reader seeks the answer in literature are always the same: are we sure to do more good than harm? With equal efficiency, can we do it cheaper? At equal cost, can we do more efficient?

The risk of developing a disease, its course and prognosis.

Risk and prognosis lie on the same continuum of disease history. The risk factor is associated with the acquisition of the disease, the “prognostic” factor with the course of the disease once acquired. The prognosis constitutes crucial information because it leads to the concept of basic risk which makes it possible to quantify the effect of an intervention and to evaluate its interest.

The determination of a causality (or etiology).

For the clinician, knowledge of etiology or causality is fundamental to his medical practice, whether prevention, diagnosis or treatment. Causality concerns the association between a risk factor and a disease, and the strength of that association.

The validity and use of a new diagnostic test.

If a reference standard exists, the article concerns the intrinsic qualities and the performance of the test. If the reference standard does not exist, the interest of studying the test lies in its clinical consequences for the patient. The question is whether the patient is better off with this diagnostic procedure than without it. This is similar to the type of article concerning the evaluation of a medical intervention in the general sense of the term.

2. What is the study plan?

Apart from the report of an interesting and unusual case, or that of a series of cases, there are four main kinds of study plan:

- **Cross-sectional study:** description of the frequency of a disease, its risk factors or its other characteristics in a given population at a given time.
- **Case-control study:** observational, retrospective study, in which the characteristics of patients with a disease (the cases) are compared with those of patients without the disease (the controls).
- **Cohort study:** observational, prospective study, in which a group of subjects exposed to risk factors for a disease is followed for a given period of time. The incidence rate of the disease in this exposed group is compared with that of a control group, followed for the same time, but not exposed to the risk factors.
- **Controlled trial:** experimental study in which an intervention is performed in a group of subjects; the outcome of this intervention is compared to that of a similar, control group that does not receive the intervention.

The reader must recognize the study plan to check if it is the most appropriate for the question asked.

Furthermore, there is a "hierarchy" among these models and the level of evidence of the results of a study (and the confidence of the reader) is variable from one model to another. It increases, from the case or the series of cases to the cross-sectional study, then to the case-control study, to the cohort study, to be maximal with the controlled trial.

3. What is the factor studied?

The factor studied is the exposure or the intervention thought to have consequences for a health problem, disease or clinical condition.

The reader should be able to know how the factor(s) was measured, whether all relevant factors were taken into account, and whether the same method of measurement was applied to all subjects, as well as from one group to other group. He must also be able to assess the quality of this measurement (variability, "blind" measurement, etc.).

If the factor studied is a diagnostic test, is there an independent comparison with the standard?

4. What is the judgment criterion?

The judgment criterion is the event or the situation supposed to be the result of the influence of the studied factor (death, illness, discomfort, dissatisfaction, etc.). The reader must find the same information as for the factor studied (precise definition, method of measurement, etc.).

5. What is the population studied and what sample is it?

The reference population, or population to be studied, is the group to which the results of the study, if valid, will apply. The sample is a subgroup of the population studied, selected, randomly or not, to represent the entire population studied, when it is not possible for practical reasons to study the latter in its wholeness.

Is the selection correct? Is there randomization? Do the groups differ in characteristics other than the factors studied? What is the proportion of subjects reaching the end of follow-up? If the factor studied is a diagnostic test, has a wide range of patients been taken into account?

It is also a question, at this stage, of judging the external validity of the study: can the conclusions, admitting that they are valid, be applied to a larger population than the simple sample studied?

6. Are there biases and confounding factors?

A bias is a systematic error that contributes to producing estimates that are systematically higher or lower than the true value of the parameters to be estimated. It is involved, for example, in the selection of patients, or in the measurement of the parameters studied.

A confounding factor is a factor that modifies the effects of the studied factor on the judgment criterion, because of its link both with the studied factor and with the judgment criterion.

Are they all considered and taken into account? If this is not the case, the validity of the study may be questioned.

7. What are the results?

Confidence in the result: the confidence interval

Clinical research is carried out most of the time on samples, for obvious reasons of feasibility. The published results are the values observed in the sample. The true value (of the treatment effect, for example), which corresponds to the truth in the population, lies somewhere around this observed value. The role of statistics is to define by calculation an interval of values where the true value is found 95 times out of 100, thus allowing the passage of the sample to the general population. This set of values constitutes the 95% confidence interval.

The reader must know that the truth lies between the two limits of the interval. The smaller this interval, the better the truth is identified, because the closer to the truth is the observed value. And this interval is all the smaller as the sample studied is large and as the number of events studied is large.

The p value

The reader must demystify the famous value of p. The value of p is the probability, calculated by the statistical test constructed from the data collected, of obtaining by pure chance a difference greater than or equal to that which is observed.

The accepted risk of asserting that there is a difference between the two groups, when in reality there is not, is called the **threshold of significance**. A significance level of 5% means that we have decided to take a risk of 5 chances out of 100 of being wrong by asserting that there is a difference.

The significance thresholds usually chosen are 5% and 1%, which means that there are only 5 chances or 1 chance in 100 that the difference observed is due to chance alone and not to the factor studied.

Statistical or clinical significance?

We can be fooled by the magic of this little p, and believe, for example, that a p value < 00001 is better than a p value < 0.05, and that this is the definitive argument to validate and accept the results.

A statistically significant difference is not necessarily clinically relevant. This is the problem of clinical significance and statistical significance. Small differences may be statistically significant if observed in large samples, but may be of little clinical importance. If a very strong association exists between a factor studied and the endpoint chosen, a small sample is sufficient to demonstrate it. On the contrary, if this association exists but is of low amplitude (10% increase in survival at 10 years, for example), then a very large sample is needed to demonstrate it.

This approach should allow the reader to keep his temper when "it is statistically significant", and not to lose hope when "it is not statistically significant".

The study power

A statistically significant difference is only of interest when it is clinically relevant.

When the difference is not significant, the reader will say that the result is negative. But is it really a “true negative”? It can indeed be a “false negative”. The difference or effect sought may well exist, but the sample size was insufficient to highlight it. Either there really is a difference or an effect, but less important than the hypothesis would have it, and here again the size of the sample was insufficient to highlight it (we speak of a lack of power of the study) .

The type II error is to assert that there is no difference or effect when in fact there is. This is the error β, and the power of the study is $1 - \beta$.

We can draw a parallel between this situation and that of false negative diagnostic tests.

8. Synthesis of the critical reading

At each of the preceding stages, was the validity of the study threatened, seriously or weakly (internal validity)? What are the authors' conclusions? Do they answer questions? Are the results applicable to the study population (external validity)?

Above all, are the results acceptable for the reader's own practice? Will they change his behavior and improve the condition of his patients?

This approach requires the reader to be familiar with the basic notions of methodology; however, it is essentially a clinical approach. Through a critical reading, the reader dismantles and evaluates the protocol that the authors have designed to carry out the study presented in the publication.

CONCLUSION

Criticism should be neither systematic nor paranoid. Although clinical research is based on rigorous rules, its aim is to analyze and quantify biological and human phenomena, the investigation of which may find its limits for methodological issues (necessary follow-up that is too long for cohorts that are too large, expected effect too weak…), budget or ethics.

Perfection does not exist. The use of this technique of literature analysis is an essentially pragmatic approach, which gives the reader the means to believe, or not to believe, and to apply, or not to apply, what could be useful to him.

This approach requires knowledge of the basic notions of methodology. It has the merit of adapting to the main types of publications. It makes it possible to assess the level of evidence provided by the authors. If there is no absolute proof, there are certainly articles on the same subject that are more convincing than others. It is up to the reader to find the best arguments with the tool offered to him!

References

American Medical Association. Guyatt GH, Rennie D. Users' Guides to the Medical Literature. WADA, 2002.

ANAES Guide to literature analysis and grading of recommendations. ANAES, Paris, 2000, www.anaes.fr

Bazi R. Documentary research. In: Matillon Y, Maisonneuve H, ed. Medical evaluation. From concept to practice. 3rd ed. Paris, Flammarion Medicine-Sciences, 2007, p.143-157.

Glanville J, Wilson P, Richardson R. Accessing the online evidence: a guide to key sources of research information on clinical and cost effectiveness. Qual Saf Health Care, 2003, 12: 229-231.

Greenhaglh T. Know how to read a medical article to decide. Medicine based on levels of evidence (evidence-based medicine) in everyday life. Translated from English by Broclain D, Doubovetzky J. Editions RanD, Meudon, 2000.

Haynes RB, McKibbon KA, Fitzgerald D et al. How to keep up with the medical literature: I. Why try to keep up and how to get started. Ann Intern Med, 1986, 105: 149-153.

Huguier M, Flahaut A. Daily biostatistics. Elsevier, Paris, 2000.

Junod AF. Medical decision or the quest for the explicit. Editions Medicine and Hygiene, Geneva, 2003.

Landrivon G. Global method of critical reading of medical articles for the use of the student and the practitioner. Frison Roche, Paris, 2002 and 2009.

Lorette G, Grenier B. Reading medical articles, Doin, Paris, 2002. Mouillet E. Bibliographic research in medicine and public health. Access guide. Paris, Elsevier, 2005.

Royle P, Waugh N. Literature searching for clinical and cost-effectiveness studies used in health technology assessments reports carried out for the National Institute for Clinical Excellence appraisal system. Health Technol Assess, 2003, 7: 1-64.

Weightman AL, Williamson J. The value and impact of information provided through library services for patient care: a systematic review. Health Info Libr J, 2005, 22: 4-25.

Appendix 1:
The reading grid

Does the information exist for each of the 8 questions?	Is the approach to the question correct?	If not, does this threaten the validity of the study?
1 - Objective - prognosis - evolution - diagnostic test - impact of an intervention - etiology - causality	Is there a hypothesis?	
2 - Type of study - case report - case series - cross-sectional study - case-control study - cohort study - controlled trial	Is the type of study appropriate to the question asked?	If not, are the results of the study completely useless?
3 - Factor(s) studied - exposure - intervention - diagnostic test	Are they well described? How are they measured? - Same measurement method for all subjects? in all groups? - Blind method? Is there an independent comparison with the reference standard?	If not, does this measurement bias threaten the validity of the study? - Idem: If not, does this bias threaten the validity of the study?
4- Judgment criterion(ies)	How are they measured? - Same measurement method for all subjects? in all groups? - Blind method? All judgment criteria relevant assessed?	- Otherwise, does this measurement bias threatens the validity of the study? - If not, do the ones that have been overlooked matter?
5 - Source population and subjects studied	- Is the selection correct? - Is there randomization? - Do the groups differ in characteristics other than the factors studied? -What is the proportion of subjects reaching the end of follow-up? - Is there a wide range of patients for the test?	If not, does this bias threaten external validity? - If not, does this bias threaten internal validity? - If not optimal, is internal validity threatened? - If not, does this bias threaten external validity?
6- Potentials Confounding factors and biases	- Are they all considered? - Are they well controlled ?	If not, does this invalidate the study?

7- Statistical analyzes and results - Confidence interval? - Statistical test? -if positive results -if negative results Strength of association Calculation of likelihood ratios	- Sufficient sample size? - Clinically relevant? - Test power, sample size?	- If not, are the results useless? - If not, is the study useful? - If insufficient, is the study useful or inconclusive?
8- Conclusions of the authors? - Answers to questions ? -Verification of the hypothesis? - Goal achieved ?	- Do the conclusions meet the objective?	Finally: - Are the results acceptable when applied to the source population? = VALIDITY - Are the results acceptable for your own practice? =APPLICABILITY

Appendix 2:

Example of use of the critical reading grid about a fictional article "Alcohol consumption and risk of breast cancer"

ABSTRACT

A case-control study was performed to determine whether alcohol consumption increases the risk of breast cancer. We interviewed 1594 women aged 22 to 56, with a recent diagnosis of breast cancer, and 1663 women of the same age, randomly selected from the general population. Women who drank alcohol did not have a higher risk of developing breast cancer compared to women who did not drink alcohol: relative risk: 1.0; 95% confidence interval: 0.8 to 1.2. Breast cancer risk was not associated with the average amount of alcohol consumed per week or the type of alcoholic beverages consumed. Compared with women who did not drink, the relative risks of developing breast cancer for women who drank beer, wine or spirits were 1.0, 0.8 and 0.9, respectively.

INTRODUCTION

Hutchinson and Bergounian's study had suggested that women who drink alcohol have a 1.5 to 2 times greater risk of developing breast cancer than women who never drink alcohol; this increased risk was associated with the consumption of all types of alcohol (beer, wine and spirits).

Breast cancer is a leading cause of death in most industrialized countries, and alcohol consumption is very common among women in these countries. If Hutchinson and Bergounian's findings—two times the risk of developing breast cancer in women who drink alcohol—apply to American women, 60% of whom drink alcohol and 7% develop cancer of the breast, then we can estimate that a non-negligible proportion of breast cancers can be attributed to alcohol consumption. It is therefore important to clarify the relationship between alcohol consumption and breast cancer.

SUBJECTS AND METHODS

The subjects participating in the study come from 8 geographical areas (the urban areas of Zorgrad, Zorgburg, Zorgcity and Zorgtown in the state of Zorgland, and the 4 urban counties of Zorgshire).

A pre-tested standard questionnaire was distributed to the women taking part in the study, at home. The questionnaire insisted on gynecological-obstetrical history and contraceptive history, family history, medical history, personal characteristics and habits, and collected information concerning the quantity and frequency of consumption of beer, wine and spirits. for the past 5 years.

The inclusion criteria for the cases were:

- age: 22 to 56 years old,
- primary breast cancer, histologically confirmed, diagnosed between January 1, 1991 and April 30, 1992,
- residing in one of the 8 zones described above.

In addition, women had to be available for questioning.

We thus included 1594 women (83.7% of women with breast cancer who met the inclusion criteria). The reasons for non-inclusion were the illness (3.4%), the patient's refusal (3.2%), the attending physician's refusal (2.9%), and the impossibility of contacting or having a interview within 6 months of the date of diagnosis (6.8%).

The controls were women identified by the *Schprountz Telephone Selection Method* living in the same geographical areas as the cases. About 94% of households have a telephone and the samples taken by random telephone calls are representative of the population. An appropriate proportion of controls by 5-year age groups was selected to be matched with breast cancer cases respecting the age distribution. Of the witnesses selected and available for questioning, 1663 (84.9%) women agreed to participate; 10.5% of the selected controls refused to participate and 4.6% had changed their place of residence or could not be contacted.

The women were asked if they had had the opportunity to drink any alcoholic beverage, or specifically beer, wine or spirits in the previous 5 years. Women who answered no were considered non-drinkers. Women who answered yes were asked the average number of days per week they drank beer, wine or spirits, and the amount they usually drank on those days. For each woman, we used the data on the amount and frequency of drinking to estimate the average number of drinks they had each week, and we multiplied this average number by 12.6 (the weight in grams of the amount absolute ethanol per drink) to estimate the weekly ingestion of pure ethanol for each woman.

We estimated the risk relative by the Cornfield method, and its 95% confidence interval by the Miettinen test.

The following variables were retained as potential confounding factors because they constitute classic risk factors for breast cancer or because they are strongly linked to alcohol consumption:

- history of benign breast disease,
- family history of breast cancer,
- age at first full-term pregnancy,
- menopausal status,
- educational level,
- age at diagnosis of breast cancer or age at questioning,
- religion,
- number of cigarettes consumed,
- and Quetelet index (weight/height2, measurement of adiposity).

We did not include the use of oral contraceptives because it has recently been proven that these are not a risk factor for breast cancer.

Logistic regression was used to simultaneously control for all of these potential confounders and to calculate the relative risk estimate for the association between alcohol consumption and breast cancer risk.

RESULTS

The age and race distribution was the same for women with breast cancer (cases) and controls. There were more nulliparas in cases than in controls, breast cancer cases were older at the birth of their first child and had more family history of breast cancer, as well as personal history of benign disease breast. A larger percentage of cases were in the premenopausal period while a larger percentage of controls had undergone surgical menopause.

Compared to those who did not drink, women who drank alcoholic beverages had a relative risk of developing breast cancer of 1.1 (95% confidence interval: 0.9 to 1.3) **(Table 1).** No influence of average weekly alcoholic consumption on the occurrence of breast cancer has been demonstrated. Women who claimed to drink the equivalent of more than 300 grams of alcohol per week had an adjusted relative risk of developing breast cancer of just 1.1 (95% confidence interval: 0.6 to 1.8).

Neither the type of alcoholic beverages nor the quantity consumed appeared to increase the risk of developing cancer, even after adjusting for the consumption of other types of alcoholic beverages **(Table 2)**. The relative risk associated with a history of heavy beer, wine or spirits consumption was 0.8, 1.2 and 1.1, respectively.

No significant association was found between breast cancer risk and alcohol consumption for women belonging to different religious groups or in different age groups; however, in general, a lower risk was observed for younger women. A relationship between alcohol consumption and breast cancer risk was not observed whether or not there was a personal history of benign breast disease or a family history of breast cancer.

DISCUSSION

Our results agree with all those who could not confirm the increased risk of breast cancer associated with alcohol consumption described by Hutchinson and Bergounian. The results of Hutchinson and Bergounian could be explained by the inclusion in their study of subjects with other alcohol use disorders; women with ovarian cancer and endometrial cancer formed a control group in their study. Currently, we are working on the hypothesis that the increased risk of developing breast cancer observed by Hutchinson and Bergounian may be due to the protective effect of alcohol on endometrial cancer rather than its effect directly on the development of breast cancer.

In this study, Hutchinson and Bergounian had limited data to investigate a dose-response relationship between alcohol consumption and breast cancer risk. There was information on the frequency, but not on the amount of alcohol consumed. In our study, we had information on both the quantity and the frequency of alcohol consumption, and we were able to estimate the average weekly alcohol ingestion.

Hutchinson and Bergounian found increased breast cancer risk for beer, wine, and spirits, although these risk estimates are based on small numbers. We did not find an increase in the risk of breast cancer associated with the consumption of each of these types of alcoholic beverages when we adjusted for each of the main risk factors for breast cancer as well as for each of the others types of alcoholic beverages. Moreover, we did not find a dose-response relationship between the risk of breast cancer and the amount of consumption of the different types of alcoholic beverages.

It is rather unlikely that biases have occurred in our results:

The selection bias was certainly very small due to the fact that the participants in the study were included only very early after the identification of the diagnosis and in the 8 geographical areas, and due to the fact that the controls were selected from the population coming from of these same areas.

It is unlikely that poor description of alcohol consumption by study participants explains the lack of association between alcohol consumption and breast cancer risk, because both cases and controls reported levels of alcohol consumption slightly higher than those reported in national surveys.

If the critical period of exposure for the development of a breast tumor is greater than 5 years before the diagnosis of breast cancer is made, then our classification into drinking and non-drinking status based on a consumption during the previous 5 years could equate patients who consumed alcohol with patients who did not consume alcohol. This misclassification could hide a real association between alcohol consumption and breast cancer if alcohol consumption during this critical period actually increased the risk of developing breast cancer. Be that as it may, the magnitude of this misclassification is certainly not greater than 5%.

Table 1 - Risk of breast cancer according to average weekly alcohol

Consumption	Cases	Controls	Relative Risk (95%)
Never drank	286	300	1,0
drank (g/week)	1308	1363	1,1 (0,9-1,3)
<50	722	759	0,9 (0,7-1,2)
50-149	342	377	0,9 (0,7-1,2)
150-199	93	87	1,1 (0,7-1,7)
200-249	56	52	1,1 (0,7-1,9)
250-299	40	37	1,0 (0,5-1,7)
≥ 300	55	51	1,1 (0,6-1,8)

Table 2 - Risk of breast cancer by type of alcoholic beverages consumed

Average consumption (g/week)	Cases	Controls	Relative risk (95%)
Never drank beer	856	896	1,0
Drank beer	738	767	1,0 (0,9-1,2)
< 50	618	629	1,1 (0,9-1,3)
50-149	82	91	0,9 (0,6-1,3)
≥ 150	38	47	0,8 (0,4-1,3)
Never drank vine	481	456	1,0
Drank vine	1113	1207	0,8 (0,7-1,1)
< 50	841	959	0,8 (0,6-1,0)
50-149	188	184	0,9 (0,6-1,2)
≥ 150	84	64	1,2 (0,8-1,9)
Never drank spirits	507	510	1,0
Drank spirits	1087	1153	0,9 (0,7-1,2)
< 50	846	897	0,9 (0,7-1,2)
50-149	164	179	0,8 (0,6-1,2)
≥ 150	77	77	1,1 (0,7-1,7)

Critical reading:

The objective of this study is to provide information regarding etiology - causation.

The hypothesis is that of the association between alcohol consumption and the development of breast cancer. Attention, "association" does not mean "cause and effect relationship".

The type of study is a case-control study. Two groups of women were formed:
- a group of breast cancer cases: 1594 women;
- a group of control women, free of breast cancer: 1663 women.

In the cases as in the controls, the investigators went back in their past to research and measure the consumption of alcohol, and to compare it between the two groups. This type of study is well suited to the question asked.

Only the retrospective model is conceivable. One can imagine the difficulty of designing a prospective study on this subject: starting from a group of "alcoholic" women and following them into the future for many years, to collect incident cases of breast cancer which would be compared to those who occur in women followed in parallel, but "non-alcoholic".

The controlled trial is of course unthinkable.

As for case series and cross-sectional studies, they would necessarily be inconclusive due to the absence of a control group.

- The factor studied (the exposure or intervention that is believed to have consequences for a health problem, disease or clinical condition) is alcohol consumption in the previous 5 years: alcohol, wine, beer, spirits . This is measured by pre-tested standard questionnaire distributed at home. We apparently avoided the risk of asking the questions differently depending on whether we are dealing with a case of breast cancer or a woman witness. This situation could indeed lead to an overestimation of alcohol consumption in cases, and therefore to an overestimation of the association between alcohol and breast cancer.

The problem lies in the quantity of alcohol ingested, in absolute terms as well as according to the different types of drink, in the two groups. The measurement method is not precise enough. It is necessary to define what a unit of wine, beer, spirits is, and to know the degree of alcohol of each drink.

- The endpoint (the event or situation assumed to be the result of the influence of the studied factor) is breast cancer. Is it the breast cancer diagnosis or is it the breast cancer mortality?

If the anatomo-pathology irrefutably defines a case, how can we be sure that a witness, in this study, is free from breast cancer? The presence of women who actually had breast cancer in the control group would lead to an underestimation of the association between alcohol and breast cancer.

Population

In the example, the reference population is that of women who consume alcohol.

In the study population, witnesses are identified by telephone. If women are not accessible by this means, or if women refuse to take part in the study, is it for socio-economic or psychological reasons, which could also explain higher than normal alcohol consumption?

In this case, does this not risk artificially increasing the rate of "alcoholics" in cases compared to controls, and going in the direction of a false association between alcohol and cancer?

The same phenomenon can occur if the cases are recruited in a hospital draining a population of a particular socio-economic level, and different from that of the controls. The 1663 female witnesses are those who agreed to respond. They represent 84.9% of identified women. Are the 15.1% who did not answer different from the others? Same question for the cases: 16.3% of identified breast cancer cases were unable to participate in the study. Are they systematically exposed in different ways to the risk factor?

Confounding factors and biases

Asking women with cancer about their alcohol consumption more carefully than controls would constitute a measurement bias. This could tend to highlight a difference between the two groups when it does not exist. This is what would also happen if the unquestionable women in the control group were consistently more alcoholics. This would be a selection bias.

Similarly, it is necessary to take into account in such a study, and this is what was done, all the other factors known to be risk factors for breast cancer (age, menopausal status, etc.).

Indeed, if they are significantly more frequent in the group of cancers, for example, we cannot know whether a possible difference in cancer rate between the two groups is due to these risk factors or to the account of alcohol consumption itself. These factors are confounding factors.

The statistical analyzes consisted of an estimation of the relative risk and its confidence interval. The intervals contain 1. So there is no association. These intervals are also very small around 1. This means that we are quite sure that this negative result is a true negative. But if it were a false negative, we would only miss an extremely weak association, an excess risk of 1.2 or a protection of 0.9, which would not necessarily be clinically relevant.

In conclusion, in the example of a case-control study on alcohol consumption and breast cancer, it is at the third stage that the internal validity of the study seems most threatened, due to the difficulty of measure the factor studied in this particular case.

With regard to the external validity of such a study, we have seen that this was threatened by the proportion and nature of the women inaccessible to the study.

Finally, is a study of this type carried out on a poor urban population in North America, or on Scandinavians, relevant for the entire French population?

Appendix 3:

Some sites

Search engines for scientific literature, online libraries:

http://www.ncbi.nlm.nih.gov/pubmed/: corresponds to the PubMed search engine site, exploring the Medline database, OldMedline, and some older indexed journals.

http://www.nlm.nih.gov/mesh/: corresponds to the PubMed database of English keywords.

http://www.crd.york.ac.uk/crdweb/: corresponds to the electronic library of the University of York (Great Britain), with in particular access to summaries from the Cochrane database and articles from health economics.

http://thomsonreuters.com/products_services/science/?view=Standard (formerly isinet.com) : complex site covering large areas of scientific publication, including the Journal of Impact Factors and Current Contents

Specific sites (thematic sites, publishing houses, examples of journals):
www.thecochranelibrary.com: website of the Cochrane Collaboration www.bdsp.tm.fr: site of the Public Health Data Bank, a public body independent of the pharmaceutical industry.

www.sciencedirect.com: brings together all the publications of the Elsevier group, i.e. approximately 2,500 periodicals.

www.bmj.com: site of the British Medical Journal

www.nejm.com: site of the New England Journal of Medicine

Institutional sites:

http://www.hon.ch/HONselect/index_f.html: site of the non-governmental organization HON, recognized by the Economic and Social Council of the United Nations. Its aim is to promote access to medical information in the broad sense of the term (including health policy).

www.has-sante.fr: site of the French High Authority for Health

http://www.ahrq.gov/: site of the 'Agency for Healthcare research and quality', depending on the US Department of Health.

Sites related to the pharmaceutical industry:

http://infodoc.inserm.fr/codecs/codecs.nsf: site of the college of health economists sponsored by INSERM (French site), largely financed by the pharmaceutical industry.
www.mdlinx.com: site selecting and summarizing articles published in peer-reviewed journals, with presentation of the summaries of the articles thus selected. Its director is an executive from Japan's leading specialty pharmaceutical marketing firm.