Pesticides increase cancer risk for farmers? Rethinking the ‘irrefutable consensus’

There is now scientific consensus that pesticides cause cancer among farmers. This conclusion has been confirmed by the courts insofar as one type of cancer, non-Hodgkin’s lymphoma, has been classified in France as a professional disease affecting farmers due to exposure to pesticides. How many farmers have fallen victim to this disease? Surprisingly INSERM (French National Institute of Health and Medical Research), when recently approached by a French inter-ministerial mission, stated that it was unable to quantify the number.

How is it that we are still at this stage for a disease whose link with agricultural workers is so well documented? To get some perspective on this, we need to look back on the history of research on farmers’ health … and how it has progressively lost traction!

Concerns stemming from retrospective case-control studies

Research into the long-term effects of pesticides on farmers took off in the 1970s as people became more aware of the effects of the accumulation of these substances in the environment. This was especially true for organochlorine pesticides. Since this was an after-the-fact analysis of the impact of products that had already been in use for a long period of time, the first epidemiological research consisted of retrospective case-control studies. This type of survey is known to be potentially susceptible to many biases but since it provides quick results it is typically used at the early stages to identify a new health problem.

These early case-control studies quickly revealed that farmers were overrepresented among the victims of a fairly large number of pathologies, mainly cancers and neurovegetative disorders. Unfortunately, these retrospective studies do not allow for an accurate calculation of the number of cases over the mean in exposed populations. Moreover, it is no easy task to reconstruct the list of products to which the farmers studied were exposed, in some cases a long time ago.

Prospective cohorts: Reassuring results

These initial disturbing results sparked a number of cohorts, i.e. studies of the health events affecting large populations of farmers still in good health and the monitoring of their use of pesticides. These cohorts are the basis for the so-called prospective studies which are much more expensive than their retrospective counterparts because they require the collection of massive amounts of information and take more than 10 years to start producing significant results.

But they are much more reliable and only they can calculate incidence rates (number of new cases each year) and mortality rates (number of deaths caused each year) among farmers and compare these figures to those of the general population. The two largest cohorts on this topic are:

- The 1993 AHS (Agricultural Health Study) cohort in the USA monitoring the health of approximately 90,000 farmers (pesticide operators and their wives)
- The larger 2005 Agrican cohort in France monitoring approximately 180,000 people, a very

significant sample of the French agricultural population. Moreover, the Agrican cohort also includes a significant proportion of farmers who do not use pesticides which could lead to the study of potentially confusing factors, i.e. health risk factors other than pesticides but which are also associated with agricultural work).

These prospective cohorts finally made it possible to calculate the standardized incidence of disease and mortality for farmers in comparison with these same figures for the general population. These studies should therefore allow researchers to identify the number of excess cases among farmers in comparison with the general population.

For example, a standardized incidence of 1.10 among farmers means that there were 10% more new cases of the disease in that group in comparison to the number of cases affecting the general population. The results of the two cohorts are very consistent … and should be reassuring judging by the traditional interpretation of standardized incidence ratios and standardized mortality ratios.

	Incidence	Mortality
Significantly above general population	Lips (US: +9% NS, **+97%S**, F: **+105%S**) Prostate (US : **+22%S**, **+16%S**, F: **+9%S**) Multiple myeloma (US : +2%NS, **+42%S**, F: **+49%S**)	None
No significant difference with general population, or discrepencies between both cohorts	Breast (US : -6%NS, -5%NS, F: +7% NS) Melanome (US : -16%NS, -4%NS, F: -2%NS) Stomach (US : -16%NS, -12%NS, F: -7 %NS) Rectum (US : **-22%S**, +10%NS, F: -10%NS) Liver & gallbladder (US : **-56%S**, +4%NS, F :**-30%S**) Testis (US : -8%NS, +10%NS, F: -31%NS) Brain & nerv. system (US : **-34%S**, -4%NS, F: +9 %NS) Hodgkin Lymphoma (US : +27%NS, NC, F: - 3 %NS) Non-Hodgkin Lymphoma (US : -5%NS, +6%NS, F: +1%NS) Leucemia (US : -6%NS, +1%NS, F : -24%NS) Anus (US : NC, NC, F: -55%NS) Thyroïd (US: -3%NS, -1%NS, F : -31%NS) Kidney (US **: -25%S**, -8%NS, F : -6%NS)	Melanome (US : -24%NS, F: -5%NS) Breast (US : -6%NS, F :+24%NS) Kidney (US : -13% NS, F :-16%NS) Hodgkin Lymphoma (US : +3%NS, F: NC) Non Hodgkin Lymphoma (US : -16% NS, F : NC) Multiple myeloma (US : +1%,NS, F : NC) Leucemia (US : -15%NS, F: NC) Thyroïd (US : +53%NS, F : NC)
Significantly lower than general population	Mouth & Pharynx (US : **-43%S**, **-45%S**, F : **-52%S**) Larynx (F : **-65%S**) Œsophagus (US : **-38S**, -32NS, F: **-43%S**) Colon (US: **-22%S**, -1%NS, F: **-18%S**) Pancreas (US : **-40%S**, -14%NS, F:**-23%S**) Respiratory tract (US : **-71%S**, **-32%S**, F: **-51%S**) Bladder (US : **-44%S**, **-37%S**, F: **-41%S**) TOTAL CANCERS (US : **-20%S** , **-8%S** , F : **-6%S**)	Lips, mouth & pharynx (US : **-66%S**, F: **-49%S**) Œsophagus (US : **-49%S**, F:**-39%S**) Stomach (US : **-48%S**, F : -16%NS) Colon (US : **-25%S**, F: **-26%S**) Rectum (US : **-31%S**, F: -12 NS) Liver & gallbladder (US : **-30%S**, F: **-31%S**) Pancreas (US : **-25%S**, F:**-22%S**) Respiratory tract (US : **-57%S**, F:**-51%S**) Prostate (US : **-19%S**, F: **-22%S**) Bladder (US : **-45%S**, F: **-41%S**) Brain & nerv. system (US : -24%S, F : NC) Lymphatique syst & blood cells, global (US : -12%NS, F: **-17%S**) TOTAL CANCERS (US : -39%S, F :-33%S)

Table 1: Standardized incidence and mortality ratios for different forms of cancer in Agrican (F) and AHS (USA) cohorts. References.

Standardized mortality of farmers is not above normal for any form of cancer. In fact, their mortality is significantly lower than that of the general population for most types of tumors while mortality is average for the rest.

Concerning the incidence of cancer, the global results for farmers are not quite as good as for mortality but are still very reassuring overall. The incidence in farmers is significantly higher for only three types of cancer: Lip and prostate cancer and multiple myeloma. We will focus on these three particular cases in a future article but for now let's start with a snapshot of the global results: for about 1/3 of the cancers studied (mainly cancers of the respiratory and digestive tracts and the bladder), the standardized incidence of cancer in farmers using pesticides is significantly lower than the mean; for nearly 2/3 of the other types of cancer, there is no significant difference with the general population.

Moreover, these non-significant results are typically less than 1 and feature relatively narrow confidence intervals making it unlikely that they hide many effects of pesticides that may have slipped through the statistical “filter.” The similarity of the Agrican and AHS studies give even greater credence to the results.

The results of these cohort studies, designed to confirm or contradict the results of the case-control studies, should therefore be considered reassuring. If one adheres to the usual interpretation of standardized incidence and mortality, there are only three forms of cancer that affect farmers significantly more than the general population and this difference is only for incidence and not mortality.

The “healthy worker effect”: A statistical truth or a pretext?

Surprisingly (or not), these results have not received much attention in official bibliographic reviews or meta-analyses such as the collective expertise of INSERM 2013, and no clear reason for these reservations has been offered. In fact, it was the aforementioned interdepartmental mission on compensation for pesticides which led INSERM to give the most transparent explanation for its reluctance: the lifestyle of farmers is associated with a number of factors protecting them from cancer.

They smoke less than the general population and often have a more balanced diet, which could explain their low incidence of respiratory and digestive tract cancer. We also see a “healthy worker” effect in other occupations involving regular physical exercise often making these professionals healthier than the average.

Consequently, INSERM warns that these favorable effects may conceal or mask the negative effects of pesticides. In other words, some pesticide-related cancers may go unnoticed when comparing farmers to the general population as the healthier lifestyle of farmers could offset this effect. This is why INSERM warns against an overly “simplistic” use of standardized incidence where normal incidence would indicate that pesticides do not have any impact.

This reluctance on the part of epidemiologists is perfectly understandable but this would mean that the standardized incidence and mortality which they continue to calculate in their publications are not valid indicators. It is therefore somewhat surprising that this problem has not yet been addressed considering that the first works on prospective cohorts date back some ten years.

It is common practice in other areas to correct incidence based on the consumption of tobacco or alcohol and a number of studies on farmers have done just that. The most general objection based on the “healthy worker” effect could be mitigated by developing standardized incidence figures where the reference population would no longer be the population at large but rather the general labor force. However, to date INSERM has not made any proposals in this regard.

We would further note that the consequence of INSERM’s reasoning should be verifiable in the Agrican cohort. In other words, if for some cancers there is a “healthy worker” effect that masks the adverse effect of pesticides, it stands to reason that the incidence of cancer in farmers who do not use pesticides should be lower than that of their pesticide using counterparts.

Indeed, the often mentioned “healthy worker effect” should be visible among both groups of farmers (non-users and conventional), the health of the non-users not being diminished by the adverse effect of pesticides. However, the most recent figures provided by Agrican on the incidence of cancer do not reflect this effect at all. In fact, they show [the opposite trend!]

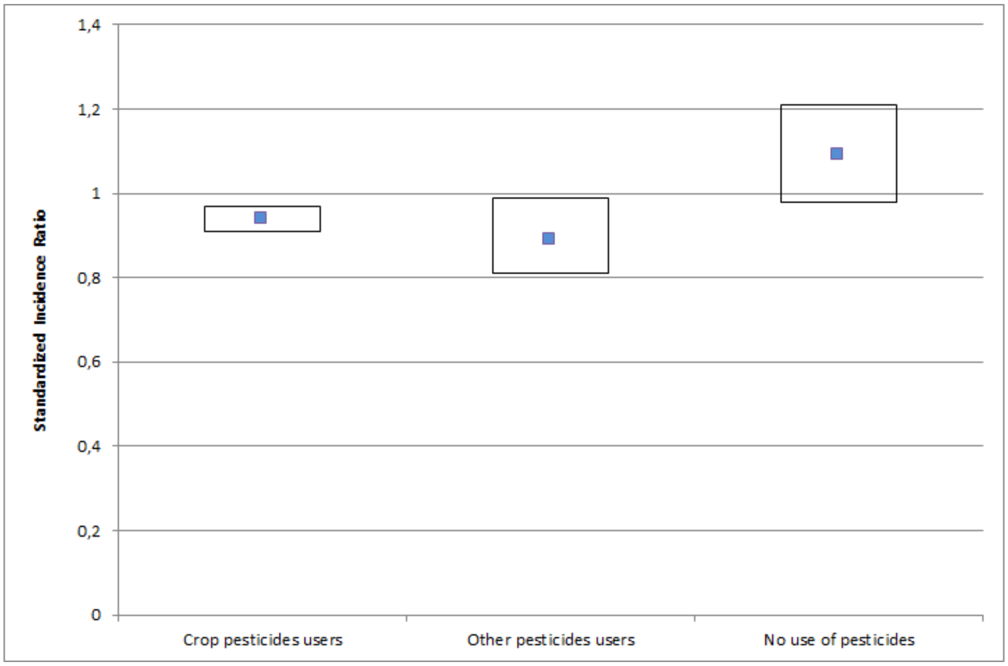

Fig 1: Standardized incidence of cancer (all types) among farmers and agricultural workers in the Agrican cohort, based on their use of pesticides (“users of other pesticides” are those who have used only veterinary products or pesticides to maintain uncultivated areas). The dots represent mean value while the vertical lines show the confidence interval (95%). The authors did not carry out a statistical analysis to check whether the differences between these three populations are significant, but we would note that the 95% confidence intervals (vertical lines) for users and non-users of pesticides do not overlap. It is therefore likely that this difference is significant … but to the detriment of the non-users of pesticides who actually suffer more cases of cancer!

Admittedly, a closer look at the different types of cancer indicates that the confidence intervals found for farmers who do not use pesticides are too high to draw reliable conclusions. But in any case, there is currently no evidence to statistically support the hypothesis that the factors protecting farmers from cancer are strong enough to mask the allegedly harmful effects of pesticides.

The rush towards irrefutability

Epidemiologists have therefore rejected the use of standardized incidence and mortality to determine whether or not pesticides cause cancer among farmers. However, the veracity of their reasonable objection (the “healthy worker” effect) has yet to be proven. Instead of trying to statistically prove the existence of the “healthy worker” effect by comparing users and non-users of pesticides to the general labor force, recent studies have opted to compare groups of farmers based on the crops they grow.

Several recent publications on the Agrican cohort highlight significant differences in the incidence of some cancers based on the crops produced or animals raised on the farm. Since pesticides are generally only used on a fairly narrow range of crops, these results are generally interpreted as indicators of the specific pesticides responsible for an excessive number of cancer cases. This is no doubt an interesting approach but there are still two rather troublesome gaps characterizing these studies:

They no longer provide a comparison with the general population. For example, it is interesting to point out that there are significantly more cases of prostate cancer among grassland farmers than among their non-grassland counterparts as a recent Agrican publication has shown. But this still does not tell us whether grassland farmers suffer more from prostate cancer than the general population.

These comparisons between agricultural production systems distinguish among a fairly large number of crops and/or animal species. In the example just cited, the authors distinguished the incidence of prostate cancer associated with 13 different crops. Analyses such as these are typically affected by the well-known “multiple comparison” effect, i.e. the risk of obtaining a statistically significant result that is simply due to chance when one conducts a large number of statistical tests. In the case of these 13 comparisons, an elementary probability calculation shows that this risk stands at 49% (1-0.9513). Results such as these should not be considered as a serious alert unless one of the following two conditions is met:

- Either its seriousness is confirmed by an additional statistical test to eliminate the “multiple comparison” effect (Bonferroni test or FDR approach), and neither of these tests were performed in this study,
- Or we observe the same type of result for the same crop and the same cancer in another cohort. To our knowledge, this link between prostate cancer and grassland farming has not been observed anywhere else. Moreover, in the Agrican cohort it is associated with an excess of prostate cancer among cattle farmers (for an obvious reason: they are more likely than other farmers to farm grasslands). However, the American AHS cohort shows no significant link between cattle farming and prostate cancer[i].

In studies such as this which compare risks based on the crop grown, significant results may be acquired but these may be due to chance and are never confirmed. They do, however, maintain suspicion on some harmful effect of the pesticides even though it is not known whether the excess number of cases observed for certain crops is due to a risk higher than that of the general population. The field of possible effects of pesticides has therefore become more and more restricted to a few types of cancer and certain crops based on scientific arguments that are increasingly ambiguous.

From scientific hypothesis to irrefutable consensus

As we have seen in this brief history of epidemiological studies, the plausibility of the possible effects of pesticides on cancer in farmers has declined steadily over time.

- The initial case-control studies suggested perceptible adverse effects for all farmers for approximately a dozen types of cancer.
- Prospective cohorts only indicate serious (but not completely consistent) evidence for three types of cancer: lip and prostate cancer and multiple myeloma. However, as we will see in a future article, recent studies on these three cancers shed very little light on the existing uncertainties, i.e. explanation for the discrepancy between incidence and mortality rates, or causes for the above average number of cases also observed in farmers who do not use pesticides.
- The current trend is to draw comparisons between subpopulations of farmers which are supposed to demonstrate the risk of pesticides associated with the crops they grow. However, the methods they use are increasingly questionable from a statistical point of view and make no comparisons with the general population.

The fact is that the amount of evidence required to suggest a carcinogenic effect of pesticides has declined steadily. Results suggesting the safety profile of pesticides (credible insofar as they confirm the validity of approval procedures) are systematically rejected and with admittedly plausible objections but ones that epidemiologists never manage to validate.

Standardized incidence rates which are normal or even less than 1 are discarded because they may be biased by a mysterious “healthy worker” effect regularly invoked by epidemiologists who never try to correct or even measure these admittedly plausible effects. The fact that there was no difference between farmers who use and those that do not use pesticides was either never addressed or was interpreted as demonstrating “contamination” of the pesticide-free group, but once again without any proof (see next article).

This change in discourse reflects the shift in the epistemological status of the hypothesis of a carcinogenic effect of pesticides in farmers. In other words, it has progressively morphed from scientific work hypothesis status to that of irrefutable consensus. However, although counterintuitive, the term irrefutable is not at all flattering. Since the work of K. Popper, it is believed that the main difference between true science and pseudo-science is its refutability: a truly scientific hypothesis is one for which one can imagine an experience that refutes it.

The initial hypothesis that pesticides cause cancer in farmers is indeed a scientific hypothesis. In other words, it can be refuted (or validated) by measuring the incidence of cancer among farmers who use pesticides and comparing that to the rest of the population. But we have seen that this hypothesis has tended more towards refutation. The new hypothesis which claims that the harmful effect of pesticides may be masked by a so-called “healthy worker” effect is also a scientific hypothesis, i.e. it can be refuted (or validated) in two ways:

- By comparing the incidence of cancer in farmers to that of other occupations involving moderate physical activity;
- Or by comparing the incidence of cancer in farmers using pesticides and farmers not using any. Since neither of these comparisons has been made to date, the refutability of this hypothesis remains entirely theoretical.

If farmer cohort studies are to regain true scientific status, it is high time that epidemiologists define what criteria they would be prepared to accept as indicating that the incidence or mortality of a type of cancer in farmers is normal. This is all the more necessary since the precautionary principle requires defining the criteria according to which a technology can be considered harmless.

This is precisely the work of the health agencies. They are the ones responsible for declaring, based on experimental results, that a product can be considered as non-hazardous. If researchers fail to decide on a set of rules allowing them to take that same decision (which is understandable after all as that is really not their job), it is incumbent upon the agencies, based on the work done by experts on the results of epidemiological studies, to draw clear and operational conclusions rather than allowing suspicion and non validated hypotheses to linger on indefinitely.

Furthermore, INSERM publications on the Agrican cohort fall prey to the major pitfall affecting the research community: focus on statistically significant results. No additional analysis is conducted on the mass of non-significant results (except meta-analyses to try to make them significant …) even though the notion of “non-significant” result actually covers two very different realities:

- Results that are truly “non-significant” because their excessive confidence interval precludes any possibility of interpretation;
- And “quasi-significant” results whose critical probability (probability that they are due to chance) is just barely over 5% and whose evidential value is therefore just below that of the ones receiving the sacred seal of “statistically significant.”

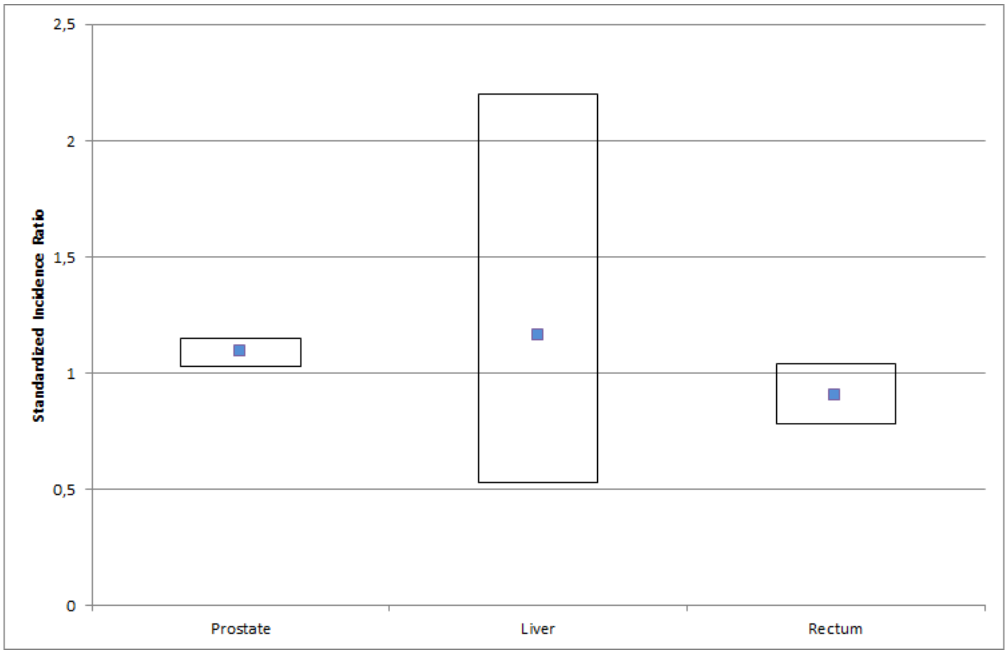

Fig 2: Examples of standardized incidence obtained for 3 forms of cancer in the most recent assessment of the Agrican cohort for farmers who use pesticides. The dots represent mean value while the vertical lines show the confidence interval (95%). In scientific publications, analysis focuses on significant results (whose confidence interval does not cross the value of 1 in red), especially if they are greater than 1 because they are the ones that indicate a link between exposure to pesticides and cancer. Very little attention is paid to non-significant results even though their value differs according to the sanitary expertise conducted. In the example given, it is impossible to conclude anything about the incidence of liver cancer because the width of its confidence interval is much greater than its deviation from 1. On the other hand, the result concerning rectal cancer, although non-significant as well, is much more relevant in sanitary expertise because the value of 1 is very close to the upper limit of its confidence interval. A simple complementary analysis would suffice to calculate the probability of an excessive number of rectal cancer cases, but it is obvious that this probability is only slightly higher than 5% (the typical significance level). Sanitary expertise needs this type of analysis highlighting the importance of non-significant results to prioritize research but researchers are not providing it.

The distinction between these two types of results will be important in guiding future research. Regarding the example illustrated in Figure 2, it is likely that uncertainty will continue to decrease regarding the results of rectal cancer as the cohort study continues to be monitored, until it reaches a level where it can be concluded, without much risk of error, that this form of cancer does not present any particular hazard.

In contrast, for liver cancer the number of cases currently observed is too low and as a result the uncertainty range is so wide that it is unlikely that clearer results will be obtained in the future from the Agrican cohort alone. Therefore, it will become vital to analyse the statistical power of Agrican's non-significant results despite the fact that this is not common practice in research (but is perfectly acceptable for health agencies).

While waiting for ANSES (The French Health Agency for Food Security) to eventually take a firm stand, we are witnessing this strange paradox: with Agrican, France has the largest prospective cohort of farmers in the world, but its decisions on the recognition of occupational diseases utterly fail to take the results of this cohort into account. None of the only three types of cancer for which Agrican provides warning signs is currently classified as a professional hazard. In contrast, non-Hodgkin's lymphoma is included in the tables of occupational diseases even though Agrican provides no evidence for this decision.

This subject alone merits a follow-up article because it is indicative of other scientific problems which are yet to be solved by the approach used by epidemiologists. This article will also provide an opportunity to examine the strength of an argument often considered decisive by opponents of pesticides: cases of dose-response relationships between exposure to pesticides and the risk of certain cancers.

Philippe Stoop is a science writer and corresponding member of the French Academy of Agriculture based in Paris, France

This article originally ran at European Scientist as Pesticides and cancer among farmers: the rush towards irrefutability and has been republished here with permission.