

**« Une passion : Connaître,
Une ambition : Transmettre »**

Recherche et évaluation des risques sanitaires : mêmes données, regards différents

Philippe STOOP

Membre Correspondant,

Section 9

Pourquoi tant de controverses?

- Toxicité du glyphosate (Roudup) :
 - Avis du CIRC en contradiction apparente avec celui des agences d'évaluation (EFSA, ANSES, ECHA, EPA,...)
- Fongicides SDHI :
 - Produits sans risques particuliers, ou scandale sanitaire futur?
- Effets sanitaires des aliments bio et « ultratransformés » :
 - Des effets négatifs des additifs alimentaires ou des résidus de pesticides, non détectés par les agences sanitaires?
- Des polémiques incompréhensibles pour le grand public, qui alimentent le discrédit des agences sanitaires

Pourquoi tant de controverses ?

➤ Des raisons médiatiques :

- Goût des medias et réseaux sociaux pour le sensationnel et la polémique
- Discrédit des « élites » et de l'entreprise

➤ Des raisons scientifico-médiatiques :

- Prise de parole croissante de chercheurs en opposition avec les agences d'évaluation :
 - CIRC (glyphosate)
 - Collectif de chercheurs (fongicides SDHI)

Mais aussi de vraies raisons scientifiques...

Pourquoi tant de controverses ?

- Un sujet abordé en 2019 par un rapport de l'OPECST * :
« Évaluation des risques sanitaires et environnementaux par les agences : trouver le chemin de la confiance »
 - Analyse des causes des incompréhensions entre citoyens et « scientifiques »
 - Propositions d'améliorations du fonctionnement des agences

- Une réflexion à prolonger :
 - Comprendre les différences de point de vue entre recherche et expertise pour y remédier
 - Tirer les conséquences de textes législatifs insuffisamment précis

* Office Parlementaire pour l'Evaluation des Choix Scientifiques et Technologiques

Recherche et évaluation des risques sanitaires : mêmes données, regards différents

➔ Les données du problème :

- Données réglementaires fournies par les industriels
- Données de recherche

➔ Retour sur deux notions de base, et leurs conséquences en évaluation des risques

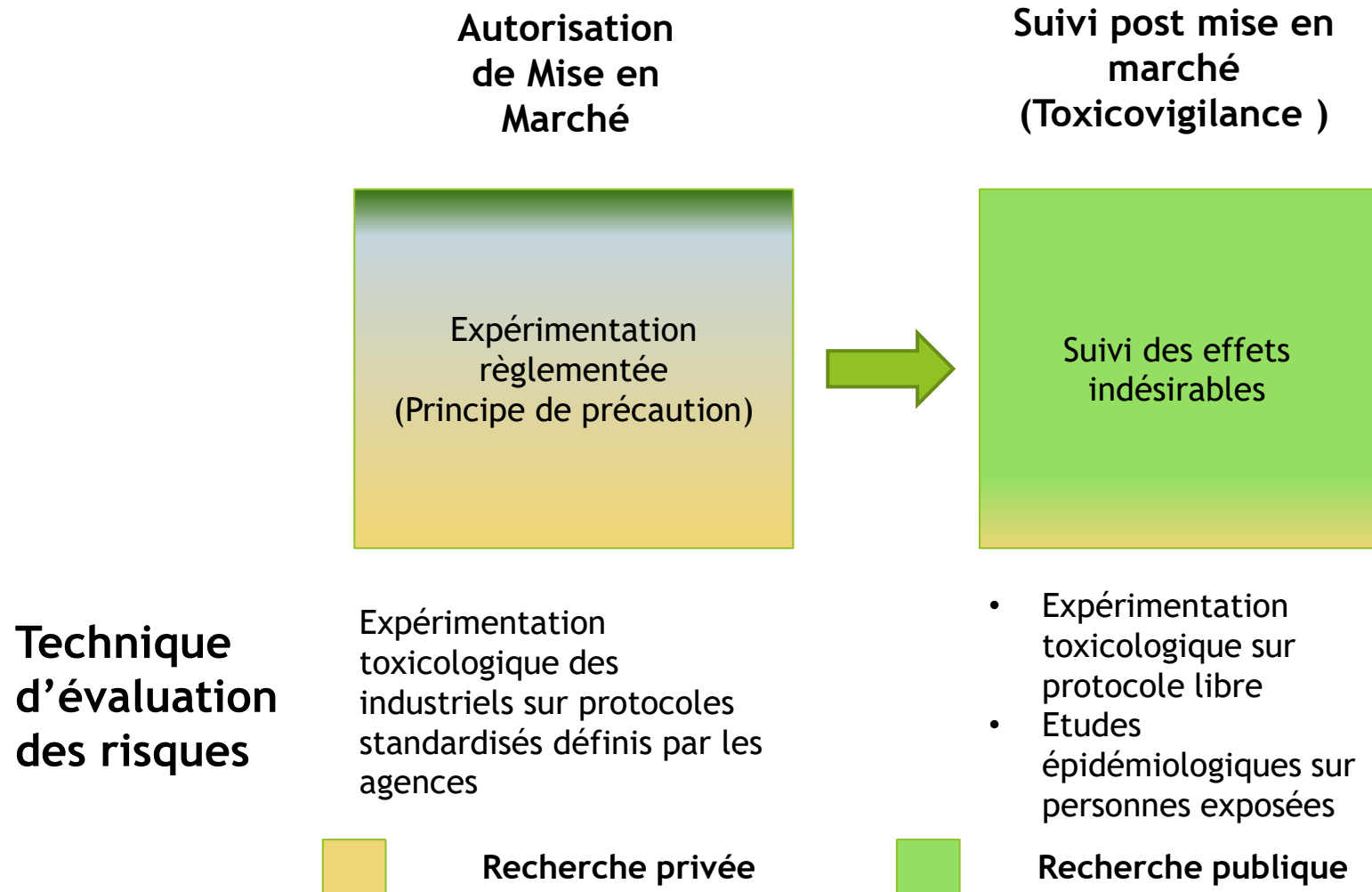
- Risques de 1ère et 2ème espèce
- Danger et Risque

➔ Les sources scientifiques de controverses

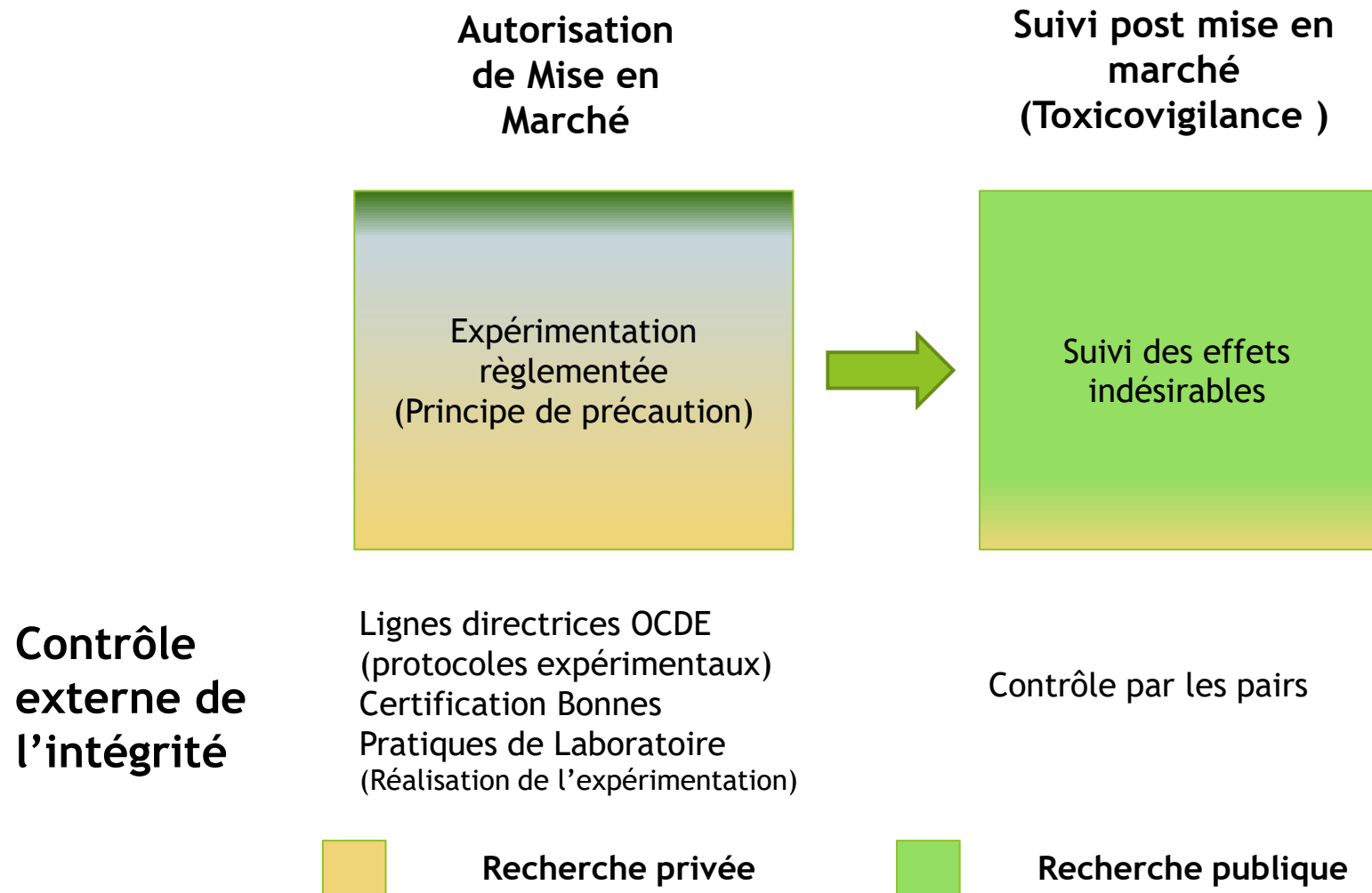
- Les tests multiples, ou le paradoxe de l’“astrologue statisticien”
- “Alphacentrisme” de la recherche et crise de la reproductibilité
- Les paradoxes de la réglementation

➔ Conclusions

Le parcours scientifique des produits règlementés



Le parcours scientifique des produits règlementés

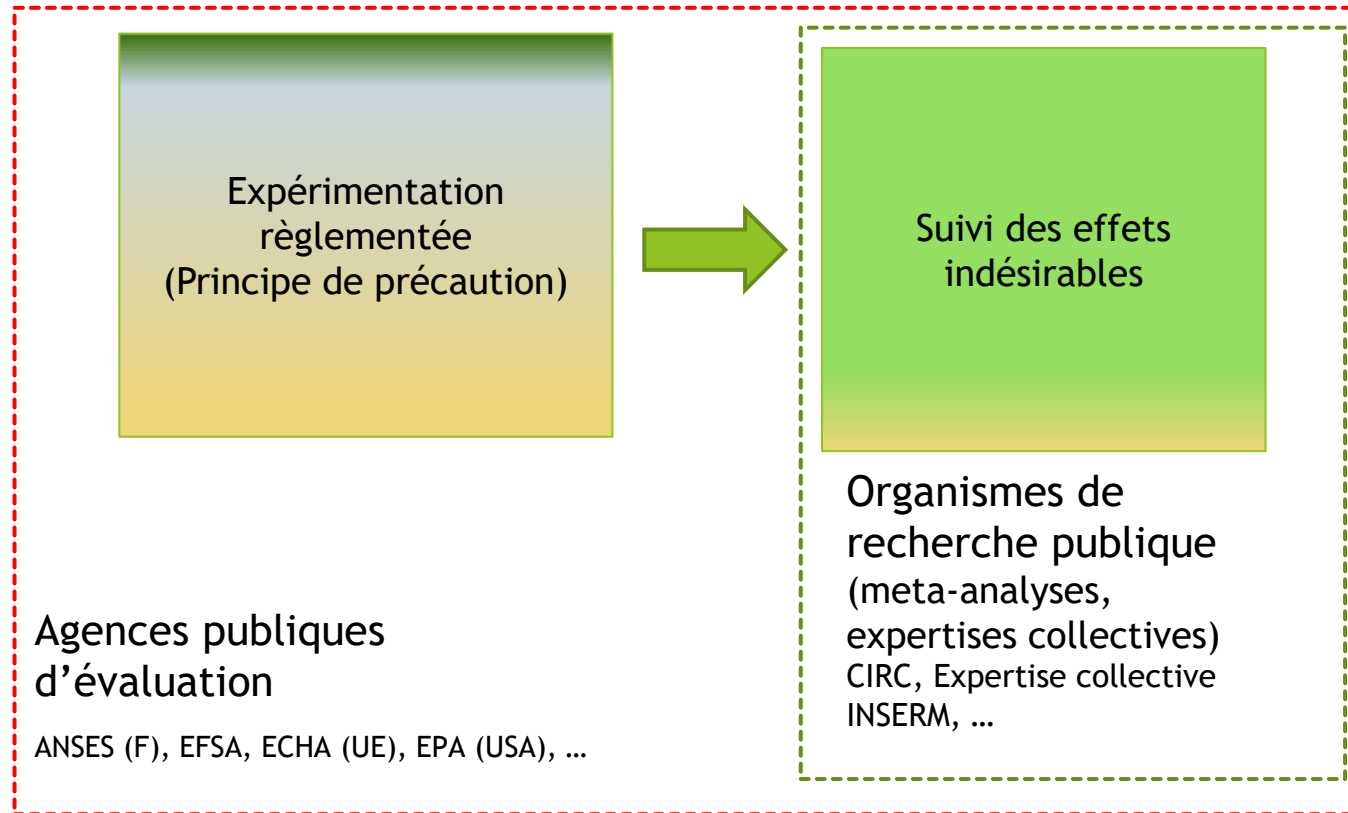


Evaluation des risques : une synthèse asymétrique

Synthèse
publique de
l'évaluation des
risques

Témoins
« à décharge »

Témoins
« à charge »



➔ Attention aux raccourcis :

- ➔ Les données confidentielles fournies par les industriels sont basées sur des protocoles définis par les autorités
- ➔ Ce ne sont pas elles qui font débat, mais les données post-homologation, issues de la recherche publique (et donc publiées)



Des scientifiques et des ONG dénoncent eux une évaluation européenne "biaisée", car basée sur des rapports fournis par les industriels tandis que, selon eux, d'autres études ne sont pas - ou mal - prises en compte.

Selon l'ONG autrichienne Global 2000, sur les 12 études disponibles sur la cancérogénicité du glyphosate, sept montrant un risque accru pour les rongeurs exposés n'ont pas été prises en compte par les agences européennes, pour des raisons de méthode statistique.

Recherche et évaluation des risques sanitaires : mêmes données, regards différents

➔ Les données du problème :

- Données réglementaires fournies par les industriels
- Données de recherche

➔ Retour sur deux notions de base, et leurs conséquences en évaluation des risques

- Risques de 1ère et 2ème espèce
- Danger et Risque

➔ Les sources scientifiques de controverses

- Les tests multiples, ou le paradoxe de l’“astrologue statisticien”
- “Alphacentrisme” de la recherche et crise de la reproductibilité
- Les paradoxes de la réglementation

➔ Conclusions

Retour sur deux notions de base, Et leurs conséquences en expertise sanitaire

- ➔ Les risques de 1^{ère} et 2^{ème} espèce α et β :
 - ➔ Risque de 1^{ère} espèce α :
Attribuer à un produit un effet sanitaire qu'il n'a pas réellement
 - ➔ Risque de 2^{ème} espèce β :
Ne pas détecter un effet sanitaire qui existe réellement
- ➔ A effectif étudié égal, les deux risques sont antagonistes
- ➔ Le risque à réduire en priorité dépend du but de l'expérimentation :
 - ➔ Pour la recherche, minimiser le risque α
 - ➔ Pour l'expérimentation d'homologation (**principe de précaution**), minimiser le risque β

Retour sur deux notions de base, Et leurs conséquences en expertise sanitaire

➔ Recherche et homologation : techniques proches, objectifs différents

	Type d'expérimentation	
	Homologation	Recherche
Protocoles	Simple et standardisés	Complexes et adaptés à chaque expérience
Risque à minimiser	2 ^{ème} espèce	1 ^{ère} espèce
Traitement des résultats incertains	Répéter expérience	Nouveau protocole pour lever les incertitudes laissées par l'expérience originale

Retour sur deux notions de base, Et leurs conséquences en expertise sanitaire

➔ Danger :

- ➔ Propriété intrinsèque des produits, des équipements, des procédés...pouvant entraîner un dommage

➔ Risque :

- ➔ Exposition d'une cible (salarié, entreprise, environnement y compris la population...) à un danger.

Le risque est caractérisé par la combinaison de la probabilité d'occurrence d'un événement redouté (accident) et de la gravité de ses conséquences.

- ➔ **Sauf exceptions, le principe de précaution est basé sur l'absence de risque, et non l'absence de danger**



Météorites :

Danger

+++++

Risque

ε

<http://www.inrs.fr/demarche/risques-industriels/definition-risque-industriel.html>

Danger / Risque :

Conséquences sur processus d'autorisation des pesticides

➔ Conséquences sur le niveau de preuve attendu :

➔ **Identification du danger :**

quelques résultats isolés, quelle que soit l'exposition, suffisent pour montrer un danger, pas de contrôle de leur cohérence

➔ **Evaluation d'un risque :**

suppose l'établissement d'une relation entre la dose d'exposition et l'effet, donc de nombreuses expérimentations, dont on peut vérifier la cohérence

<https://www.usinenouvelle.com/article/avis-d-expert-les-impasses-d-une-reglementation-des-perturbateurs-endocriniens-basee-sur-le-danger.N539619>

Recherche et évaluation des risques sanitaires : mêmes données, regards différents

➔ Les données du problème :

- Données réglementaires fournies par les industriels
- Données de recherche

➔ Retour sur deux notions de base, et leurs conséquences en évaluation des risques

- Risques de 1ère et 2ème espèce
- Danger et Risque

➔ Les sources scientifiques de controverses

- Les tests multiples, ou le paradoxe de l’“astrologue statisticien”
- “Alphacentrisme” de la recherche et crise de la reproductibilité
- Les paradoxes de la réglementation

➔ Conclusions

Le paradoxe de l' « Astrologue statisticien » (problème des test multiples)

- ➡ Dans les publications scientifiques, on considère un résultat comme statistiquement significatif si son risque de 1^{ère} espèce est inférieur à 5%.
- ➡ Un astrologue veut démontrer l'existence d'un lien entre le signe astrologique du Cancer, et le risque d'avoir un cancer : quelle probabilité $P_{5\%}$ pour qu'il obtienne un résultat statistiquement significatif ($\alpha = 5\%$)?
 - ➡ Pour le cancer en général : $P_{5\%} = 5\%$
 - ➡ S'il étudie 20 localisations de tumeurs, $P_{5\%} = 1 - 0,95^{20} = 64\%$

Le paradoxe de l' « Astrologue statisticien » (problème des test multiples)

- ⊕ Dans une expérimentation complexe, on obtient **nécessairement** des résultats significatifs, **même si le produit testé est inoffensif**
- ⊕ La question n'est plus de savoir si on a des résultats significatifs, mais si on a **un excès de résultats significatifs** par rapport au risque de 1^{ère} espèce
 - ⊕ plus de 5% de résultats significatifs
 - ⊕ plus de 1% de résultats hautement significatifs

Le paradoxe de l' « Astrologue statisticien » (problème des test multiples)

- ➔ **Comment reconnaître un résultat “vraiment” significatif?**
- ➔ Tests statistiques complémentaires pour contrôler l'effet “tests multiples” :
 - ➔ Test de Bonferroni
 - ➔ Procédures FDR (False Discovery Rate)
- ➔ Des tests rarement pratiqués... et encore plus rarement exploités!

Le paradoxe de l' « Astrologue statisticien » (problème des test multiples)

➡ Exemple d'une publication sur l'effet des perturbateurs endocriniens sur les troubles du comportement des jeunes garçons :

- ➡ Association significative entre exposition prénatale à 3 produits et troubles du comportement
- ➡ 13 produits étudiés, croisés avec 7 indicateurs du comportement, soit 91 combinaisons :
- ➡ 5,22% de résultats significatifs à 5%, 9,89% de résultats significatifs à 10% :
mêmes résultats qu'avec le signe astrologique
- ➡ Probabilité critique après test FDR : 0,42 : confirme que les résultats "significatifs" ont 42% de chances d'être dus au risque de 1^{ère} espèce (et donc 58% de chances d'être "vraiment significatifs")

➡ Conclusions des auteurs (dans le titre et dans l'abstract) :

"Several phenol and phthalate biomarkers were associated with increased scores on the SDQ subscales at 3and/or5y." (Test FDR cité seulement dans le paragraphe "Résultats").

Vu dans la presse

➔ Service de presse INSERM :

« association » entre exposition
et troubles du comportement,
causalité pas évoquée explicitement



Une étude épidémiologique menée par l'Inserm^[1] sur les familles de la cohorte EDEN (500 garçons nés entre 2003 et 2006 et leurs mères) montre que l'exposition pendant la grossesse à certains phénols et phtalates est associée à des troubles du comportement des garçons entre 3 et 5 ans. Les composés les plus préoccupants à cet égard sont le bisphénol A, le triclosan et le di-n-butyl phtalate, ou DBP. Les résultats viennent d'être publiés dans la revue *Environmental Health Perspectives*.

Vu dans la presse

➡ Dans la presse :

Lien de causalité

(au conditionnel pour les plus prudents)



Le paradoxe de l' « Astrologue statisticien » (problème des test multiples)

➔ Comment confirmer un résultat significatif?

Type de cancer	Personnes de signe astrologique Cancer									
	France		USA		Chine		Brésil		Moyenne	
	OR	p	OR	p	OR	p	OR	p	OR	p
Lèvres-cavité buccale	1.21	0.65	2.18	0.02	1.12	0.73	0.92	0.78	1.40	0.06
Poumon	0.84	0.69	2.25	0.01	1.26	0.49	0.89	0.71	1.35	0.09
Estomac	1.59	0.28	0.72	0.29	1.16	0.66	1.05	0.88	1.07	0.69
Foie	0.82	0.65	1.11	0.74	0.69	0.28	0.69	0.23	0.73	0.06
Rein	1.60	0.28	0.74	0.33	1.10	0.79	1.06	0.85	0.95	0.76
Prostate	1.19	0.69	0.62	0.13	0.86	0.64	1.07	0.82	0.98	0.90
Vessie	1.11	0.81	0.36	< 0.01	1.05	0.89	0.53	0.04	0.70	0.07
Colon	0.82	0.65	1.96	0.04	1.45	0.27	1.53	0.17	1.26	0.19
Cerveau	0.97	0.95	1.06	0.86	1.13	0.72	1.19	0.58	1.03	0.87
Lymphome de Hodgkin	0.84	0.67	0.82	0.53	0.53	0.07	0.81	0.49	0.79	0.16
Lymphome non hodgkinien	1.07	0.88	0.48	0.02	0.93	0.83	0.41	< 0.01	0.74	0.08
Testicule	1.92	0.14	0.98	0.95	0.68	0.26	0.62	0.13	0.90	0.55
Thyroïde	0.52	0.15	1.58	0.16	0.52	0.06	1.03	0.93	1.00	0.99
Pancreas	0.66	0.36	1.24	0.50	0.76	0.42	1.45	0.24	1.02	0.92
Melanome	1.06	0.89	2.02	0.03	0.60	0.13	1.53	0.18	1.27	0.18

OR : Odds Ratio p : Probabilité critique

NB : chiffres provenant d'une publication scientifique réelle, concluant à un effet de l'exposition domestique aux pesticides sur le risque de cancer (Impact factor de la revue : 13)

Interprétation astrologique :

- des corrélations significatives observées dans 2 des pays étudiés
 - des effets significatifs à 10% pour 5 types de cancers
- => Il y a des liens significatifs entre cancer et signe astrologique

Interprétation statistique :

- un excès de résultats significatifs à 5%, mais aucun au seuil de 10%
 - distribution symétrique des résultats
 - aucune cohérence entre les pays
- => Aucune démonstration d'un effet du signe astrologique

Le paradoxe de l' « Astrologue statisticien » (problème des test multiples)

⊕ Le problème de la confirmation des résultats incertains :

- ⊕ La seule façon vraiment sûre de confirmer un résultat à significativité incertaine : retrouver EXACTEMENT le même résultat sur un jeu de données indépendant du premier
- ⊕ En pratique, cette replication est rarement réalisée
- ⊕ Pour que le résultat incertain puisse être confirmé (ou infirmé), il faut qu'il ait été publié

⊕ Les publications que nous venons de citer sont légitimes, mais ne peuvent être utilisées sans précaution dans une évaluation de risque.

Recherche et évaluation des risques sanitaires : mêmes données, regards différents

➔ Les données du problème :

- Données réglementaires fournies par les industriels
- Données de recherche

➔ Retour sur deux notions de base, et leurs conséquences en évaluation des risques

- Risques de 1ère et 2ème espèce
- Danger et Risque

➔ Les sources scientifiques de controverses

- Les tests multiples, ou le paradoxe de l’“astrologue statisticien”
- “Alphacentrisme” de la recherche et crise de la reproductibilité
- Les paradoxes de la réglementation

➔ Conclusions

Idée reçue n° 1

- ⊕ Dans les revues scientifiques, le contrôle par les pairs ne garantit pas la véracité des résultats présentés. Il garantit :
 - Leur (forte) **plausibilité**
 - Leur **reproductibilité**
- ⊕ Mais les mécanismes de publication et d'évaluation de la recherche font que cette répliquabilité reste le plus souvent théorique :
 - Réplication réussie : aucun intérêt pour la presse scientifique, donc aucun pour l'évaluation du chercheur
- ⊕ Conséquence :
 - Accumulation de résultats significatifs non confirmés, sur des protocoles difficilement comparables

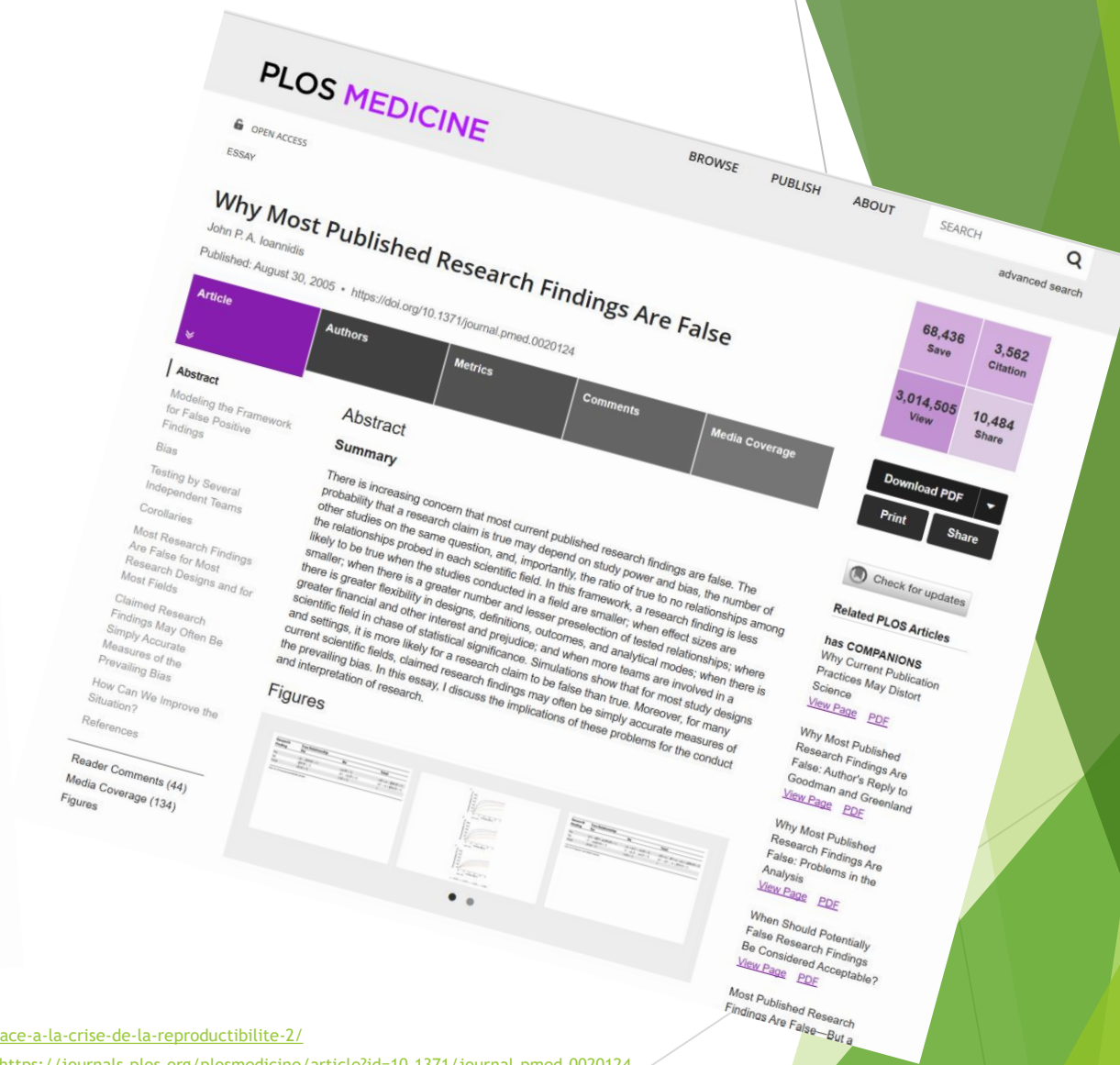
La crise de reproductibilité de la recherche

- ➡ Tentatives de replications systématiques de publications scientifiques :
% de résultats confirmés < 50%



<https://www.ouvriirlascience.fr/repenser-la-robustesse-et-la-fiabilite-en-recherche-les-chercheurs-face-a-la-crise-de-la-reproductibilite-2/>

<https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020124>



Epidémiologie : des difficultés d'interprétation

- ➔ Epidémiologie :
- ➔ *Des interactions trop complexes pour être bien redressées par les traitements statistiques*
- ➔ Difficulté d'isoler l'effet d'un facteur nutritionnel favorable ou défavorable



<https://jamanetwork.com/journals/jama/article-abstract/2698337>

Epidémiologie : des difficultés d'interprétation

➔ Dans une même cohorte, risque de cancer réduit de :

- 25% pour les forts consommateurs de bio
- 20% pour les faibles consommateurs d'aliments "ultratransformés"
- 24% pour les faibles consommateurs de viande rouge

Chaque facteur étudié est supposé redressé des autres facteurs de confusion nutritionnels

Notes Académiques de l'Académie d'agriculture de France
Academic Notes from the French Academy of Agriculture
(N3AF)
Note de recherche

Santé et alimentation : attention aux faux-semblants statistiques !

Philippe Stoop¹

¹ ITK, Cap Alpha, avenue de l'Europe, 34830 Clapiers, France.

Correspondance :
philippe.stoop@itk.fr

Résumé

Ces derniers mois, plusieurs études épidémiologiques très médiatisées déclarent avoir observé des liens statistiques forts entre le risque de cancer et la consommation d'aliments bio ou ultra-transformés. Toutefois ces résultats reposent sur des traitements statistiques complexes, dont un examen attentif montre la fragilité : prise individuellement, chacune de ces publications semble minimiser fortement l'effet des facteurs de risque de cancer bien établis. En outre, quand on les compare entre elles, on constate qu'elles se contredisent mutuellement, alors qu'elles ont été réalisées sur la même population. Bien que parus dans des revues prestigieuses, ces travaux confortent donc les inquiétudes d'épidémiologistes éminents, qui appellent à refonder les méthodes employées en épidémiologie nutritionnelle.

publication seems to underestimate the effect of well-established cancer risk factors. Moreover, when we compare them, we find that they contradict each other, although they were made on the same population. Although published in prestigious journals, these studies reinforce the concerns of eminent epidemiologists, who call for a re-founding of the methods used in nutritional epidemiology.

Mots clés

nutrition, cancer, aliments biologiques, aliments ultra-transformés, statistiques

Keywords

nutrition, cancer, organic food, ultraprocessed food, statistics

Vu dans la presse



Attention :

- aux vision manichéennes sur les divergences scientifiques à propos des résultats de la recherche
- Aux remises en cause du doute scientifique



Recherche et évaluation des risques sanitaires : mêmes données, regards différents

➔ Les données du problème :

- Données réglementaires fournies par les industriels
- Données de recherche

➔ Retour sur deux notions de base, et leurs conséquences en évaluation des risques

- Risques de 1ère et 2ème espèce
- Danger et Risque

➔ Les sources scientifiques de controverses

- Les tests multiples, ou le paradoxe de l’“astrologue statisticien”
- “Alphacentrisme” de la recherche et crise de la reproductibilité
- Les paradoxes de la réglementation

➔ Conclusions

Les paradoxes de la réglementation

- ➔ Une réglementation basée sur le danger et non plus sur le risque :
 - ➔ Le raisonnement « classique » basé sur le risque est critiqué pour deux raisons :
 - ➔ Pour les perturbateurs endocriniens : résultats scientifiques en faveur d'un effet dose non monotone
 - ➔ Pour les produits cancérogènes : difficulté et délai important pour la mise en évidence des effets cancérogènes chez l'homme
- ➔ Depuis 2011 (règlement CE 1107/2009), un pesticide ne peut être autorisé s'il est classé cancérogène probable ou perturbateur endocrinien probable (en terme de danger, quel que soit le risque)
- ➔ Une décision raisonnable, mais dont les conséquences réglementaires n'ont pas été tirées

Définition du danger : Des critères très souples...

- ⊕ **Critères de définition pour des « preuves suffisantes de cancérogénicité » en expérimentation animale (Règlement CE no 1272/2008) :**
 - « un lien de causalité est établi... dans (a) au moins deux espèces animales ou (b) au moins deux études indépendantes sur une espèce...[ou] une incidence accrue de tumeurs chez les deux sexes d'une même espèce dans une étude correctement réalisée, de préférence selon les bonnes pratiques de laboratoire. »
- ⊕ **Lien de causalité :**
 - « liaison significative (à 5%) entre exposition à l'agent étudié et une incidence accrue des néoplasmes malins ou d'une combinaison donnée de néoplasmes bénins et de néoplasmes malins »
- ⊕ **Des critères très souples, sans conséquences dans une réglementation basée sur le risque, mais aberrants (et non révisés) dans une réglementation basée sur le danger.**

Définition du danger : Des critères très souples...

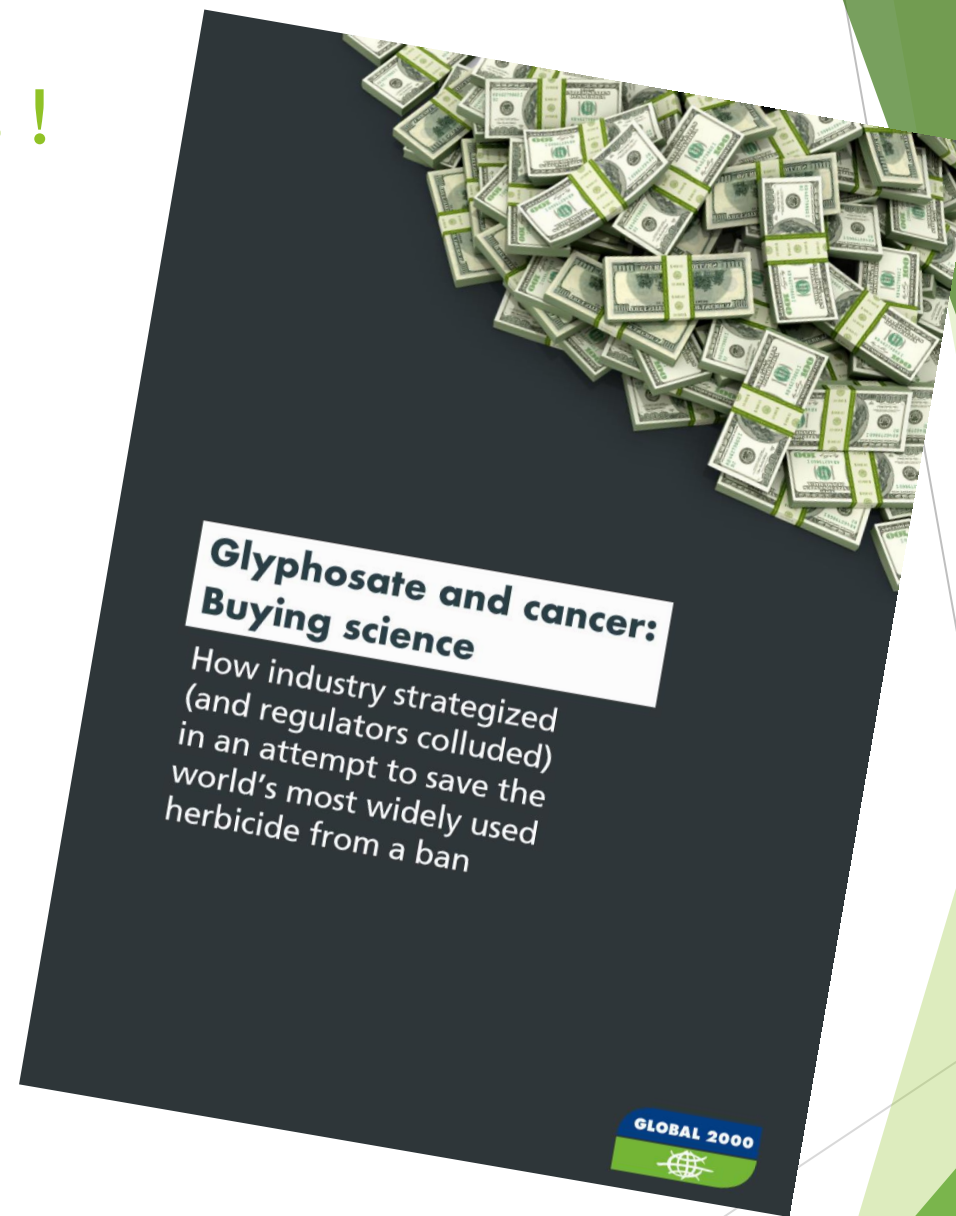
➔ Pour un produit inoffensif :

- La probabilité de trouver par erreur un effet cancérigène dans une étude est de 5% (=1/20)
- Avec 40 études (ou 40 tests statistiques différents, dans un nombre plus faible d'études), il devient très probable d'avoir 2 études indépendantes montrant une liaison significative à 5% (« un lien de causalité ») : problème des tests multiples

➔ Appliquées sans discernement, les lignes directrices officielles conduisent inévitablement à « trouver » un danger pour tout produit très étudié

Définition du danger : ...et des interprétations souples !

- Le règlement CE no 1272/2008 invoqué comme démonstration d'une collusion entre industriels et agences sanitaires (page 31)



Evaluation des risques sanitaires et environnementaux : Pourquoi tant de controverses ? - 30 mai 2018

- ➔ Les données du problème
- ➔ Retour sur deux notions de base, et leurs conséquences en expertise sanitaire
 - ➔ Danger et Risque
 - ➔ Risques de 1^{ère} et 2^{ème} espèce
- ➔ Les sources scientifiques de controverses
 - ➔ Le paradoxe de l’“astrologue statisticien”
 - ➔ La “cacophonie de la recherche”
 - ➔ Les paradoxes de la réglementation européenne
- ➔ En guise de conclusion

Conclusions

1. Recherche et évaluation des risques : deux approches différentes, toutes deux légitimes

- ➡ Il est normal que les revues scientifiques acceptent des publications de recherche qui restent à confirmer
- ➡ Il est normal que les agences sanitaires rejettent certaines de ces publications, si elles ne sont pas conformes à leurs critères de qualité (qui sont différents de ceux de la recherche)

2. Ces différences de méthode doivent être assumées et expliquées clairement aux citoyens pour restaurer la confiance

- ➔ Exemple : ne pas laisser croire que l'avis du CIRC sur le glyphosate est compatible avec celui des agences d'évaluation sanitaire
 - Sur l'animal, le CIRC a retenu des expérimentations non conformes aux lignes directrices OCDE, donc rejetées par toutes les agences
 - Sur l'homme, le CIRC a retenu des études épidémiologiques chez les agriculteurs, donc en exposition réaliste, donc il considère bien qu'il y a risque et pas seulement danger.

3. Le législateur doit préciser clairement ses attentes en matière de niveau de preuve scientifique

- ➔ Respect des lignes directrices OCDE : obligatoire ou facultatif? Si facultatif, quelles limites ?
- ➔ Faut-il adapter le règlement CE 1272/2008 pour tenir compte de la situation nouvelle créée par le CE 1107/2009?

Idée reçue n°2

Non, la rigueur scientifique ne nuit pas à la santé !

- ⊕ Respect des lignes directrices OCDE : garantit la puissance statistique des expérimentations (détection des « signaux faibles »)
- ⊕ L'EFSA prend des décisions de classement plus restrictives en moyenne que le CIRC





Questions ?

Philippe STOOP

Membre Correspondant
Section 9

