

L'efficacité du dépistage mammographique est-elle vraiment contestable?

Une revue critique de l'article de Gøtzsche et Olsen

C. Bouchardy^a, L. Raymond^a, F. Levi^b, N. Probst-Hensch^c,
M. Tubiana^d, C. de Wolf^e

- ^a Institut de médecine sociale et préventive, Genève
^b Institut universitaire de médecine sociale et préventive, Lausanne
^c Institut für Sozial- und Präventivmedizin der Universität Basel, Bâle
^d Centre Antoine-Béclère, Faculté de Médecine, Paris, France
^e Zentralinstitut für die kassenärztliche Versorgung in der Bundesrepublik Deutschland, Cologne, Allemagne

Résumé

Gøtzsche et Olsen ont publié dans *The Lancet* du 8 janvier 2000 [1] un article contestant l'efficacité de la mammographie de dépistage. Ils ont critiqué la qualité méthodologique des études prospectives randomisées publiées. Ils ont jugé non optimale la qualité de la randomisation des études ayant montré une baisse de la mortalité par cancer du sein. Ils en déduisent que le dépistage du cancer du sein par la mammographie est inefficace et donc injustifié.

La Ligue suisse contre le cancer, dans le cadre de laquelle sont élaborés les programmes nationaux de lutte contre le cancer, a demandé une expertise épidémiologique de cet article. En tenant compte des éléments de réponse disponibles dans la littérature et après examen approfondi des arguments avancés par Gøtzsche et Olsen, nous concluons que l'efficacité du dépistage du cancer du sein par mammographie ne peut pas être remise en cause par cet article. En effet, l'exercice statistique auquel se sont livrés Gøtzsche et Olsen présente des défauts méthodologiques ne permettant pas des conclusions valides sur l'efficacité de la mammographie.

Correspondance:
Dr Christine Bouchardy
Centre médical universitaire
Institut de médecine sociale et préventive
CH-1211 Genève 4

Introduction

Un article publié par Gøtzsche et Olsen dans *The Lancet* en janvier 2000 [1] met en doute l'efficacité même de la mammographie: les auteurs critiquent la qualité méthodologique des études randomisées sur lesquelles reposent les principales preuves de l'efficacité du dépistage. Gøtzsche et Olsen affirment que les études «positives» (celles qui ont montré une baisse significative de la mortalité) reposaient sur une randomisation non «optimale», contrairement aux études négatives: ils en tirent la conclusion que le dépistage est inefficace et donc injustifié. Les arguments et les conclusions de cet article ont été immédiatement réfutés par de nombreux professionnels [2–11].

La Ligue suisse contre le cancer, dans le cadre de laquelle sont élaborés les programmes nationaux de lutte contre le cancer, a souhaité disposer d'une expertise épidémiologique de l'article de Gøtzsche et Olsen. Le présent rapport a pour objectif de revoir de façon critique les arguments avancés par Gøtzsche et Olsen, en rappelant notamment les arguments déjà évoqués à ce propos dans la littérature, et de discuter du bien fondé des conclusions de ces auteurs.

L'article de Gøtzsche et Olsen

Suite à la publication d'un article [12] concluant à l'absence de baisse de la mortalité par cancer du sein en Suède, où le dépistage par cancer du sein par la mammographie est recommandé depuis 1985, Gøtzsche et Olsen ont revu les 8 études randomisées sur l'efficacité du dépistage du cancer du sein par la mammographie: Malmö, Canada, Göteborg, Stockholm, Deux comtés (traités séparément), New York et Edimbourg. Les auteurs se sont focalisés sur trois sources de biais potentiels, classiquement discutés dans les essais thérapeutiques, à savoir: a) une randomisation non «optimale»; b) l'absence de décision à l'aveugle pour juger de la survenue de l'événement déterminant l'efficacité finale du dépistage (à savoir les décès par cancer du sein); et c) les exclusions après randomisation. Parallèlement, les auteurs ont cherché à déterminer la morbidité entraînée par le dépistage. Leurs principales conclusions commentées plus loin peuvent être résumées comme suit:

- Les études ayant montré une baisse de mortalité par cancer du sein, à savoir Göteborg, Stockholm, Deux comtés, New York, Edimbourg et la phase II de l'étude de Malmö sont jugées biaisées par Gøtzsche et Olsen, notamment en raison de différence d'âge ou de couche sociale entre le groupe des femmes invitées et leurs témoins. Ces biais pourraient expliquer l'excès de mortalité générale qui avait été observé pour l'ensemble des études suédoises.

- Les études n'ayant montré aucune baisse de la mortalité par cancer du sein, à savoir l'étude de Malmö et l'étude canadienne, sont jugées par Gøtzsche et Olsen non biaisées sur la base des mêmes critères. C'est en fonction des résultats de ces deux études que le dépistage mammographique est considéré comme inefficace, donc injustifié.

Critiques scientifiques spécifiques

Etude de Sjönell et Ståhle

Il est important de rappeler que l'analyse de Sjönell et Ståhle [12], qui est à l'origine de l'étude de Gøtzsche et Olsen a été très critiquée et considérée comme imparfaite, notamment en termes de puissance, d'interprétation des effets temporels et d'une mauvaise identification des cohortes dépistées [9, 13–19].

Etude suédoise des Deux comtés

L'étude suédoise des Deux comtés est la cible des critiques les plus vives de la part de Gøtzsche et Olsen. C'est la plus grande des études randomisées et elle influence de façon importante les résultats des méta-analyses auxquelles elle est intégrée. La baisse de la mortalité de 30% entre groupes invité et non invité y persistait, même après 20 ans suivant la randomisation [20]. Les critiques de Gøtzsche et Olsen sont les suivantes:

- le groupe invité est en moyenne de 5 mois plus âgé que le groupe témoin (dans une population de 133 000 femmes âgées de 40 à 74 ans); les modalités de la randomisation sont tenues par Gøtzsche et Olsen pour responsables de cette différence d'âge;
- l'analyse des résultats bruts révèle un léger excès de décès toutes causes dans le groupe invité, dû à la différence d'âge mentionnée plus haut;
- le nombre de femmes incluses dans l'étude présente des différences d'une publication à l'autre.

Ces critiques avaient déjà fait l'objet d'une réponse en 1989 de la part des investigateurs, qui avaient commenté ces différences d'âge et reconnu que le groupe invité était en moyenne légèrement plus âgé que le groupe témoin, ce qui expliquait la faible (et non significative) élévation du taux de mortalité dû à d'autres causes [21]. Cette différence d'âge a été d'ailleurs corrigée par un ajustement adéquat par les auteurs [21] confirmant une baisse importante de la mortalité par cancer du sein (risque relatif non ajusté: 0,70; risque relatif ajusté: 0,69) et une disparition complète de l'augmentation de la mortalité (risque relatif non ajusté: 1,02; risque relatif ajusté: 0,99). De plus, comme de Koning l'a soulevé dans son commentaire sur l'article de Gøtzsche et Olsen paru simultanément, cette différence d'âge minime peut être considérée comme «conservatrice», à savoir qu'elle entraîne une sous-estimation des bénéfices du dépistage [3].

L'inconsistance des effectifs rapportés avait aussi déjà fait l'objet d'une explication antérieure par les

investigateurs de l'étude suédoise des Deux comtés [21]. Dans l'analyse originale, publiée en 1985 [22], les femmes chez lesquelles un cancer du sein avait été diagnostiqué avant l'étude avaient été exclues des cas et des décès par cancer du sein dans le groupe invité et le groupe témoin. Par la suite, une fusion avec les données des registres des tumeurs a permis d'étendre cette exclusion à l'ensemble des femmes randomisées. Ainsi, depuis 1989, les études portent sur cette population «épurée», qui compte quelques centaines de femmes de moins que la population initiale.

Par ailleurs, et même s'il est vrai que la technique de la randomisation par grappes reste moins efficace que la randomisation individuelle, celle-ci est toutefois judicieuse et tout à fait valide, voire préférable, lorsque la randomisation porte sur des groupes de femmes résidant dans des zones économiquement très diversifiées [23].

Etude de Göteborg

Gøtzsche et Olsen ont noté une différence d'un mois (jugée significative) en ce qui concerne l'âge moyen des femmes des groupes invité versus témoin, dans les résultats publiés chez les femmes âgées de 39 à 49 ans [24], ce qui, à leurs yeux, justifie son exclusion des méta-analyses. En l'occurrence, la randomisation a été effectuée par cohorte d'années de naissances successives. Le coefficient de randomisation ayant varié en fonction de la capacité à prendre en charge le dépistage, les différences d'âge étaient attendues et l'analyse d'ajustement sur l'âge avait été planifiée a priori et non pas a posteriori, comme affirmé par Gøtzsche et Olsen. Idéalement, l'ajustement aurait dû se faire sur l'année de naissance et non sur l'âge. Une telle analyse montre un risque relatif de mortalité par cancer du sein de 0,54 (IC 95%: 0,30–0,96), identique à celui qui est obtenu par l'analyse ajustée selon l'âge, à savoir 0,55 (IC 95%: 0,31–0,96) [25].

Etude de Malmö

Gøtzsche et Olsen n'ont retenu que les premiers résultats négatifs de l'étude de Malmö (Malmö I) et exclu les résultats publiés ultérieurement avec un recrutement plus important et un suivi plus long (Malmö II), lesquels montraient en revanche un risque relatif de mortalité par cancer du sein de 0,81. Gøtzsche et Olsen justifient cette démarche par le fait que les derniers résultats de Malmö mélangeaient l'ensemble des cohortes de femmes incluses et que suite à une erreur administrative, les femmes nées en 1934 n'avaient pas été randomisées mais toutes incluses dans le groupe dépisté. Comme les auteurs ne possédaient pas de distribution par âge, ils ont estimé de façon très compliquée et approximative l'âge moyen du groupe invité de l'étude Malmö à partir des données de la méta-analyse des études suédoises: «We used the number of women as reported in the meta-analysis and the mean ages as estimated above. We took account of the fact that women in Göteborg were randomly allocated to study and control groups in the approximate ratio of 1.2 in the 39–49-year age-group and 1.6 in the 50–59-year age-group. We had

Tableau 1

Les principales différences entre les études d'intervention en population et les essais thérapeutiques.

Essais thérapeutiques	Etudes d'intervention en population
<i>Exemple</i>	
Efficacité de la prise d'une molécule supposée active versus placebo	Efficacité de l'invitation systématique à la mammographie versus pratique médicale courante
<i>Type de sujet concerné</i>	
Malade	Population en théorie saine
<i>Nombre de sujets concernés</i>	
Dizaines ou centaines, rarement des milliers	Milliers ou dizaine de milliers
<i>Comparabilité des groupes</i>	
Des différences sur les critères pronostiques ne sont pas acceptables	Des différences même minimales deviennent significatives en raison du grand nombre de sujets impliqués
<i>Compliance dans le groupe traité/d'intervention</i>	
Importante ou totale	Partielle
<i>Ecart au protocole/Contamination du groupe témoin</i>	
Nul ou extrêmement faible	Inévitable et parfois important, devant être documenté
<i>Perdu de vue ou exclusions après randomisation</i>	
Exceptionnel	Inévitable

no data on age for the 50–59-year group, but since the imbalance in age in the 39–49-year group was numerically small, we used a mean age of 54 for both study and control groups. For Malmö, we used 57 years as estimated mean age in the study group, similar to the Kopparberg and Östergötland trials. This approach yielded a mean age in the study groups of 54.93 years, very close to the 55.05 years reported in the meta-analysis. Since the mean age in the control groups was 0.51 years lower, that in the Malmö control group was estimated to be 56.85 years. The difference of 0.15 years is not significant ($z = 1.53$, $p = 0.13$) which suggests that the randomisation method in Malmö was adequate» [1]. Les auteurs en ont déduit que la différence d'âge, chiffrée à 0,15 ans et basée sur des groupes comprenant chacun 21 000 femmes n'était pas significative. A noter que pour l'étude de Göteborg, la différence d'âge qui n'était que de 0,09 ans et portait sur 12 000 femmes dans chaque groupe avait été jugée significative.

L'étude canadienne

L'investigation des données de l'étude canadienne reste curieusement superficielle, au regard de la vigueur de la critique des études positives. Gøtzsche et Olsen acceptent sans réserve l'adéquation de la randomisation de l'étude canadienne, malgré l'évidence de biais incomparablement plus inquiétants que des différences d'âge, à savoir notamment la haute prévalence de tumeurs avancées à l'entrée dans l'étude

et le fait que cette étude puisse à peine être considérée comme une étude de population étant donné qu'un quart seulement des femmes invitées à participer a accepté d'être inclus dans l'étude [26–33]. De plus, il est étonnant que Gøtzsche et Olsen affirment que l'étude canadienne est de loin la mieux documentée, tout en affirmant qu'ils n'ont retrouvé aucune information sur la distribution par âge des groupes concernés. Ce n'est qu'a posteriori, après parution de leur article, qu'ils ont obtenu cette information [34].

Critiques scientifiques d'ordre général

Les défauts méthodologiques

Gøtzsche et Olsen n'ont pas eu accès à l'ensemble des données individuelles des études randomisées et ont dû se contenter des données publiées souvent incomplètes, déjà regroupées ou éparées. De façon peu systématique, semble-t-il, ils ont pris contact avec les responsables des études afin d'obtenir plus de détails notamment sur les méthodes de tirage au sort des femmes. A les lire, ils n'ont pu obtenir que très partiellement les informations souhaitées. D'après eux, les données sur l'âge manquaient dans 2 études sur 8 et les caractéristiques sociales dans 5 études sur 8. Leur revue aurait donc dû être d'emblée reconnue comme non réalisable. Par la suite, les auteurs ont paru reconnaître ce fait [35]. Précisons que dans plusieurs cas, les auteurs ont pallié ce manque d'information par des extrapolations pour le moins hasardeuses.

Contrairement à ce qu'affirme Gøtzsche et Olsen, les différences d'âge ne sont pas forcément le fait d'une randomisation non optimale. Les lois statistiques nous apprennent que même des différences dont la probabilité de survenue est rare peuvent se produire. Il est manifeste que Gøtzsche et Olsen n'ont pas l'habitude de conduire des études sur des groupes de population générale, pour lesquelles le tirage au sort est effectué usuellement à partir de listes de personnes résidentes. Ces listes ne sont jamais parfaitement à jour et comprennent toujours un certain nombre de personnes qui ont quitté la région (émigration ou réémigration). Or, ce n'est que parmi les sujets qui sont l'objet de l'étude que l'on est informé de ces départs (par exemple par le retour postal des lettres d'invitation à participer). Il s'ensuit inévitablement un certain nombre de différences entre le groupe avec intervention et le groupe témoin, notamment en ce qui concerne l'âge, l'ampleur de la différence dépendant alors de la structure par âge du mouvement d'émigration. Un ajustement a posteriori est de ce fait souvent effectué pour annuler ces différences.

En tout état de cause, les auteurs auraient dû rappeler que les études d'intervention en population ne sont assimilables qu'en partie aux essais thérapeutiques. Les spécificités des deux objectifs, telles que celles décrites dans le Tableau 1 auraient impérativement dû être évoquées. Il aurait notamment fallu tenir compte du fait que le très grand nombre de femmes impliquées dans ce type d'étude tend à rendre signifi-

ficatives des différences d'âge même négligeables (de quelques mois), alors que dans des essais thérapeutiques portant sur une centaine de sujets des différences de plusieurs années peuvent ne pas l'être. La comparabilité entre les groupes ne peut pas toujours être maintenue dans les études de population, en raison notamment des inévitables perdus de vue; il est alors d'usage de procéder aux ajustements nécessaires, alors que cela n'est pas justifié dans les études thérapeutiques. Les taux d'incidence ou de mortalité non standardisés selon l'âge rendent toute comparaison régionale ou internationale impossible. Gøtzsche et Olsen contestent pourtant fermement cette approche courante en épidémiologie. Ils raisonnent en se plaçant dans le cadre usuel des essais cliniques, lesquels visent à mesurer l'efficacité réelle du traitement, alors qu'un effet de dilution (p.ex. non-participation des femmes invitées et participation des femmes du groupe témoin) est inévitables dans les études d'intervention en population. En outre, on rappellera qu'il est normal de considérer qu'un seul essai clinique est concluant, alors que les conclusions des études d'intervention en population sont en général moins définitives, demandant à être répétées pour que soit établie la causalité entre l'intervention et l'effet.

Gøtzsche et Olsen ont souligné à juste titre que le reproche que l'on peut faire à certaines études randomisées est de ne pas avoir déterminé la cause de décès à l'aveugle. Ce problème avait déjà été décrit et discuté (voir notamment [36]). D'ailleurs, lors de la méta-analyse des études suédoises, les décès ont été revus en masquant le statut de dépistage [37]. De plus, on doit rappeler que le problème ne se pose pas pour les études cas-témoins, puisque celles-ci partent de la cause de décès pour remonter au dépistage; or ces études ont montré une baisse importante de la mortalité par cancer du sein chez les femmes participantes.

La méthode statistique visant à évaluer l'hétérogénéité entre études est insuffisamment décrite. Les auteurs ont utilisé un «Fix effect model», à savoir un modèle fixe dont les variables sont définies à l'avance et qui ne tient compte que des facteurs de confusion définis a priori sans permettre de tester leurs interactions. Les variables introduites dans le modèle n'ont pas été décrites. Alors que le seuil de signification traditionnellement utilisé dans les études épidémiologiques ou cliniques est fixé à 5%, Gøtzsche et Olsen ont fixé dans certains cas ce seuil à 10% ce qui leur permet de décrire une hétérogénéité significative entre études retenues et exclues ($p = 0,08$), hétérogénéité qui n'aurait pas été constatée au seuil conventionnel de signification. La méthode d'étude des différences entre groupe invité et groupe témoin, notamment en ce qui concerne l'âge n'est pas décrite. On ne comprend pas si l'âge à l'entrée dans les études a été considéré en années et mois (sous forme de variable continue), en années arrondies ou encore sous forme de groupes d'âge. Selon l'approche retenue, l'estimation des âges moyens et des écarts types s'en trouve inévitablement modifiée. Les estimations des

auteurs ne peuvent donc qu'être très approximativement reconstituées. Il reste ainsi surprenant que de petites différences d'âge (p.ex. 0,09 dans l'étude de Göteborg) soient jugées significatives alors que ne le sont pas des différences plus marquées impliquant un nombre de femmes plus important (p.ex. 0,15 dans l'étude de Malmö). Pour l'étude de Göteborg, en particulier, nos propres calculs n'ont d'ailleurs pas confirmé l'existence d'une différence statistiquement significative.

Enfin, si l'on reconnaît que les inévitables biais potentiels inhérents à chacune des études tendent à minimiser ou au contraire augmenter l'estimation de l'efficacité réelle du dépistage, une méta-analyse incluant l'ensemble des études est au contraire indispensable. On peut ainsi douter de conclusions ne reposant que sur 20% des données provenant des études randomisées. Il est évident qu'une omission de cette nature dans une méta-analyse ou une revue de la littérature portant sur l'efficacité d'une thérapeutique aurait été accueillie avec scepticisme.

Gøtzsche et Olsen ne discutent pas l'ampleur des biais introduits sur l'estimation de l'efficacité de la mammographie. Ils se contentent de rappeler qu'en théorie, dans les essais cliniques, des biais de randomisation peuvent expliquer la majeure partie, voire la totalité de l'efficacité thérapeutique mise en évidence. Dans le cas particulier, il est clair que les différences d'âge relevées entre le groupe d'étude et le groupe témoin ne peuvent pas expliquer une baisse de la mortalité de l'ordre de 25%. D'après les calculs auxquels nous nous sommes livrés à partir des données de mortalité suédoises sur le cancer du sein de 1995 (données OMS), un écart d'âge moyen d'un an entre deux groupes de population compris entre 40 et 70 ans ne pourrait expliquer une différence relative de mortalité par cancer du sein que de l'ordre de 6%. Or, la différence d'âge la plus importante relevée par Gøtzsche et Olsen (étude Kopparberg) ne se chiffre qu'à 0,45 année d'âge, d'où une différence de mortalité de moins de 3%. De plus, Gøtzsche et Olsen ne discutent pas le fait que des études avec des biais potentiels opposés convergent dans leurs résultats. En effet, dans l'étude suédoise des Deux comtés les femmes du groupe invité étaient plus âgées (de 0,45 ans pour Kopparberg et 0,27 ans pour Östergötland), alors que dans l'étude de Stockholm les femmes invitées étaient plus jeunes que les témoins (de 0,18 ans). Ces différences d'âge devraient entraîner une sous-estimation de l'efficacité du dépistage dans le premier cas et une surestimation dans le second. Ils n'ont pas tenu compte non plus des différences de l'efficacité en fonction de l'âge au dépistage. Il n'y a finalement aucune évidence que la randomisation ait systématiquement induit des biais en faveur du groupe invité. De plus, si la baisse de la mortalité par cancer du sein observée dans les études suédoises n'était due qu'à ce biais, il n'y aurait pas eu de raison qu'elle n'ait pas été aussi marquée avant 50 ans, âge pour lequel il est difficile de contester une moindre efficacité de dépistage [38-40]. De même, le fait que les femmes invitées appartiendraient à une

Tableau 2

Les principaux commentaires et conclusions de Gøtzsche et Olsen sur les études randomisées et le risque relatif de mortalité par cancer du sein.

Etude	Conclusions émises par Gøtzsche et Olsen	Principaux commentaires sur les différences observées entre groupes invité et témoin	Risque relatif de mortalité par cancer du sein
Malmö I	Adéquate	Distribution par âge non disponible Différence calculée par approximation à 0,15 an*, jugée non significative	0,96
Canada	Adéquate	Distribution de l'éducation comparable Distribution par âge non disponible**	1,08
Malmö II	Inadéquate	Distribution par âge non disponible Femmes nées en 1934 non randomisées mais toutes incluses dans le groupe invité	0,81
Göteborg	Inadéquate	Différence d'âge jugée significative: femmes plus jeunes de 0,09 an dans le groupe invité	0,55 (age: 39-49)
Stockholm	Inadéquate	Différence d'âge jugée significative: femmes plus jeunes de 0,18 an dans le groupe témoin	0,73
Deux-comtés (Kopparberg et Östergötland)	Inadéquates	Différences d'âge jugées significatives: femmes plus âgées de 0,45 an et de 0,27 an dans le groupe témoin, respectivement pour Kopparberg et Östergötland	0,70
New York	Inadéquate	Différences de facteurs de risque, notamment l'éducation	0,79
Edinburgh	Inadéquate	Différence dans la couche sociale	0,87

* age approximatif estimé par Gøtzsche et Olsen à partir des données de la méta-analyse des études suédoises.

** age non retrouvé par Gøtzsche et Olsen.

couche sociale plus favorisée (comme c'est apparemment le cas dans l'étude new-yorkaise et celle d'Edimbourg) aurait normalement entraîné une sous-estimation de l'effet du dépistage, puisque le risque de cancer du sein est plus élevé dans les milieux favorisés.

Un manque d'impartialité évident

Aucune étude d'intervention portant sur plusieurs centaines ou milliers d'individus ne peut être parfaite. Des écarts aux protocoles sont inévitables. On pourrait premièrement contester le fait que l'existence de la différence d'âge constitue le critère dichotomique et définitif du biais de randomisation. En effet, il est infondé d'attribuer exclusivement à la randomisation des différences d'âge, quel que soit le nombre de sujets. De plus, sur des effectifs aussi énormes, des différences d'âge minimales deviennent significatives. Sur 8 études, le jeu des probabilités veut qu'au moins une d'entre elles présente des différences significatives. Et encore, ce critère n'est-il pas appliqué de manière rigoureuse puisque l'étude canadienne et la première phase de l'étude Malmö ont été acceptées malgré l'absence de description de l'âge, alors que les études suédoises étaient exclues sur la base de différences minimales, mais jugées significatives (tab. 2).

Gøtzsche connaît parfaitement le biais lié au jugement des investigateurs lorsque celui-ci n'est pas effectué à l'aveugle. Il en avait décrit les dangers dans un article quelque peu provocateur «Blinding during data analysis and writing of manuscripts» [41]. Leur

choix du critère de l'adéquation de la randomisation et l'intensité avec laquelle ils recherchent les biais potentiels apparaissent plutôt dictés par le résultat des études que par des erreurs de randomisation évidentes et objectives. Comme par hasard, l'adoption de ce critère de sélection amène Gøtzsche et Olsen à rejeter toutes les études dont les résultats sont en faveur du dépistage. Le manque d'impartialité est d'autant plus probable que l'on connaît le contexte dans lequel le rapport de Gøtzsche et Olsen a été initié et les a priori fortement négatifs du premier auteur face au dépistage en général «Do we wish to turn the world's healthy citizens into fearful patients-to-be who, in the not too distant future, might be asked to deliver, for example, annual samples of faeces, urine, sputum, vaginal smear, and blood, and undergo X-ray and ultrasound examination with all they entails in term of psychological morbidity and the potential for harm because of further testing and interventions due to false positive findings?» [42]. Cet a priori négatif face au dépistage aurait nécessité une analyse critique des études sans connaissance de leurs résultats.

L'étude n'est pas une méta-analyse de l'Organisation Cochrane

Gøtzsche et Olsen donnent une fausse impression que leur article du *Lancet* constitue un rapport officiel de l'Organisation de recherche internationale collaborative Cochrane, laquelle s'est publiquement distanciée de cette publication. La réponse de Andy Oxman (Chairperson of Cochrane Collaboration Steering

Group) est claire à ce propos: «Although Gøtzsche et Olsen have a Cochrane protocol in The Cochrane Library, they have not yet submitted a Cochrane review to the Cochrane Breast Cancer Group. What they published in the Lancet is not a Cochrane review and has no Cochrane status other than arising from two people who work in a Cochrane Centre. The Cochrane Collaboration had no control (editorial or otherwise) over this review» [43, 44]. L'idée d'éditer un rapport officiel du Cochrane Steering Committee on breast cancer est actuellement considérée (www.cochrane.org).

Conclusion

On pourrait résumer le cheminement du raisonnement de Gøtzsche et Olsen de la façon suivante: parce qu'il existe des différences d'âge entre le groupe des femmes invitées et non invitées, nous suspectons que la randomisation de ces études n'est pas «optimale» et nous ne devons pas en tenir compte. Si nous excluons les résultats de ces études, alors nous n'avons plus de preuve que le dépistage est efficace et celui-ci n'est par conséquent pas justifié.

A partir d'une analyse épidémiologique, comme on l'a vu imparfaite, Gøtzsche et Olsen tirent des conclusions qui se veulent définitives et qui sont en fait totalement disproportionnées. Du point de vue scientifique, il est déraisonnable de remettre en question l'efficacité du dépistage du cancer du sein par la mammographie suite à l'article de Gøtzsche et Olsen. L'exercice statistique de ceux-ci, s'il reste stimulant, n'est ni rigoureux, ni objectif. Il ne répond pas aux exigences que l'on est en droit d'attendre des études évaluant une mesure de cette portée pratique. Fortement critiqué par les experts internationaux en matière de dépistage du cancer du sein, il est à remettre dans le contexte d'un débat marginal des quelques opposants au dépistage, à la recherche de publicité.

Références

- Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable. *Lancet* 2000;355:129-34.
- Cates C, Senn S. Screening mammography re-evaluated (letter). *Lancet* 2000;355:750.
- De Koning HJ. Assessment of nationwide cancer-screening programmes. *Lancet* 2000;355:80-1.
- De Landtsheer JP, Delaloye JF, Hessler C, De Grandi P, Paccaud F, Levi F. Dépistage organisé du cancer du sein. L'expérience vaudoise. *Rev Med Suisse Romande* 2000; 120:501-10.
- Duffy SW, Tabàr L. Screening mammography re-evaluated (letter). *Lancet* 2000;355:747-8.
- Hayes C, Fitzpatrick P, Daly L, Buttner J. Screening mammography re-evaluated (letter). *Lancet* 2000;355:749.
- Law M, Hackshaw A, Wald N. Screening mammography re-evaluated (letter). *Lancet* 2000;355:750.
- Moss S, Blanks R, Quinn MJ. Screening mammography re-evaluated (letter). *Lancet* 2000;355:748.
- Nyström L. Screening mammography re-evaluated (letter). *Lancet* 2000;355:748-9.
- Palmieri C, Fishpool S. Breast cancer screening. *Br Med J* 2000;321:567-8.
- Wald N. Populist instead of professional. *J Med Screen* 2000;7:1.
- Sjönell G, Ståhle L. Hälsokontroller med mammografi minskar inte dödlighet i bröstcancer. *Läkartidningen* 1999; 96:904-13.
- Rosen M, Lundin A, Nyström L, Rutqvist LE, Stenbeck M, Talbäck M. Incidence and mortality of breast cancer during a 25-year period. International and regional comparisons. *Läkartidningen* 2000;97(4):294-9.
- Rosen M, Nyström L, Stenbeck M. The National Board of Health and Welfare: stronger evidence basis for the benefits of mammography. *Läkartidningen* 2000;97(7):739-40.
- Rosen M, Stenbeck M. The debate on mammography, the National Board of Health and Welfare is answering Sjönell and Ståhle. *Läkartidningen* 2000;97(8):859-60.
- Rehnqvist N, Rosén M, Karlberg I. The National Board of Health and Welfare on mammographic screening: mortality analysis requires completely different methods. *Läkartidningen* 1999;96(9):1050-1.
- Rosen M, Rehnqvist N. Breast screening. No need to reconsider breast screening programme on basis of results from defective study. *Br Med J* 1999;318(7186):809-10.
- Rutqvist LE. Naturalförloppet, grova metoder ledde till felkalkyl om bröstcancer. *Läkartidningen* 1999;96:1210-11.
- Tabàr L. Assessment of the effect of mammographic screening can not be based on wrong data. *Läkartidningen* 1999;96(14):1763-4.
- Tabàr L, Vitak B, Chen HH, Duffy SW, Yen MF, Chiang CF, et al. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am* 2000;38(4):625-51.
- Tabàr L, Fagerberg G, Duffy SW, Day NE. The Swedish two county trial of mammographic screening for breast cancer: recent results and calculation of benefit. *J Epidemiol Community Health* 1989;43:107-14.
- Tabàr L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Gröntoft O, et al. Reduction in mortality from breast cancer after mass screening with mammography: randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. *Lancet* 1985;i:829-32.
- Duffy SW, South MC, Day NE. Cluster randomization in large public health trials: the importance of antecedent data. *Stat Med* 1992;11(3):307-16.
- Bjurstam N, Björnelid L, Duffy SW, Smith TC, Cahlin E, Eriksson O, et al. The Gothenburg breast screening trial. First results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization. *Cancer* 1997;80(11):2091-9.
- Bjurstam N, Björnelid L, Duffy SW, Prevost TC. Author reply. *Cancer* 1998;83:188-90.
- Kopans DB. Canadian National Breast Screening Study. *Lancet* 1997;350(9080): 810.
- Tarone RE. The excess of patients with advanced breast cancer in young women screened with mammography in the Canadian National Breast Screening Study. *Cancer* 1995; 75(4):997-1003.
- Baines CJ. A different view on what is known about breast screening and the Canadian National Breast Screening Study. *Cancer* 1994;74(4):1207-11.
- Kopans DB, Halpern E, Hulka CA. Statistical power in breast cancer screening trials and mortality reduction among women 40-49 years of age with particular emphasis on the National Breast Screening Study of Canada. *Cancer* 1994; 74(4):1196-203.
- Kopans DB, Halpern E, Hulka CA. Mammography screening for breast cancer. Reply to the commentaries. *Cancer* 1994; 74(4):1212-6.
- Mettlin CJ, Smart CR. The Canadian National Breast Screening Study: an appraisal and implications for early detection policy. *Cancer* 1993;72:1461-5.
- Miller AB, Baines CJ, To T, Wall C. Canadian national breast screening study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. *Can Med Assoc J* 1992; 147(10):1459-76.

- 33 Miller AB, Baines CJ, To T, Wall C. Canadian national breast screening study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. *Can Med Assoc J* 1992; 147(10):1477-88.
- 34 Miller AB, Baines CJ, To T, Wall C. Screening mammography re-evaluated (letter) *Lancet* 2000;355:747.
- 35 Gøtzsche PC. Openness about the Swedish breast cancer screening trials is needed. *Läkartidningen*. 2000;97(3):2-3.
- 36 Hurley SF, Kaldor JM. The benefits and risks of mammographic screening for breast cancer. *Epidemiol Rev* 1992; 14:101-30.
- 37 Nyström L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. *Lancet* 1993; 341(8851):973-8.
- 38 De Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. *J Natl Cancer Inst* 1995;87(16):1217-23.
- 39 Tabár L, Fagerberg G, Chen HH, Duffy SW, Smart CR, Gad A, Smith RA. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. *Cancer* 1995;75(10):2507-17.
- 40 Tabár L, Fagerberg G, Duffy SW, Day NE, Gad A, Gröntoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am* 1992;30:187-210.
- 41 Gøtzsche PC. Blinding during data analysis and writing of manuscripts. *Control Clin Trials* 1996;17:285-93.
- 42 Gøtzsche PC. Screening for colorectal cancer. *Lancet* 1997; 349(9048):356. Discussion 358.
- 43 Dean PB. The Swedish mammography screening trials. Check up on your sources. *Läkartidningen* 2000;97(3):2.
- 44 Dean PB. The articles by Gøtzsche and Olsen are not official Cochrane review and lack scientific merit. *Läkartidningen* 2000;97(3):3.