

35^{èmes} Journées des Économistes de la Santé Français,
5 et 6 décembre 2013, Créteil

L'impact causal de la survenue du cancer sur le parcours professionnel : l'apport des données administratives

T. Barnay (Upec, Erudite), M. Ben Halima (Irdes),
E. Duguet (Upec, Erudite), J. Lafranchi (Paris II, Cee),
C. Le Clainche (Cee, Ens)

Discutant : M Audibert, CERDI-CNRS, UdA



Contexte de l'étude

L'étude se situe dans un contexte français i) de forte augmentation de l'incidence des cancers, ii) de dépistage précoce, iii) de l'allongement de la durée de cotisation pour la retraite.

- Une part des cancers survient lorsque la personne est en activité
- La survenue d'un cancer devrait avoir des répercussions sur la réinsertion professionnelle de la personne.

La revue de la littérature sur le lien entre cancers et emploi montre que :

- la littérature internationale est abondante,
- l'impact négatif sur la trajectoire professionnelle dépend des limitations fonctionnelles ;

- les taux d'emploi des personnes ayant survécu à leurs cancers est de 7 à 8 points (en %) plus faibles que ceux des salariés en bonne santé ;
 - la littérature s'est intéressée principalement aux individus âgés de plus de 50 ans,
 - les travaux en France sont peu nombreux.
- Intérêt de mener des études en France, en s'intéressant aux effets différenciés selon le genre et au-delà de deux années (5ans)

Objectif de l'étude

- Estimer l'impact de la survenue du cancer sur la carrière (les parcours professionnels) à CT et MT selon le genre, l'âge et les profils de carrière précédant la survenue du cancer.

Données : étude rétrospective basée sur des données administratives, la base HYGIE (2005-2008) qui regroupe des données de la CNAV (assurance vieillesse) et de la CNAM (caisse d'assurance maladie des salariés - retraités et salariés du secteur privé).

Ces données sont recueillies à partir du système national de gestion des carrières et sont ensuite appariées aux données de prestations d'assurance maladie, tirées du système d'information inter-régime de l'assurance maladie. Elle permet d'identifier les affections de longue durée, dont les tumeurs malignes (code : ALD30).

Avantage de HYGIE : base nationale avec une couverture importante (cotisants du régime général de l'assurance maladie) et des données détaillées (âge, genre, consommation médicale, arrêt de travail, etc.).

Inconvénient : tous les cas de cancer ne sont pas inscrits en ALD30

→ Différence (mais non importante) de mesure de l'incidence des cancers estimée à partir de cette base ou des registres des cancers

Caractéristiques des individus base Hygie (2005-2008) :

- 528.496 salariés secteur privé, entre 22 et 70 ans (en 2005), 47,6% de femmes. 79.933 souffrent d'une affection de longue durée, dont 13.981 avec une ALD 30 (cancer), soit 17,5% de l'échantillon et 55% sont des femmes.

Stratégie économétrique

Méthode de la double différence avec appariement des individus avec et sans cancer pour estimer l'effet causal du cancer sur le parcours professionnel des malades.

Cette méthode nécessite :

i) la constitution de deux groupes, un groupe avec intervention (appelé traité) et un groupe sans intervention (appelé témoin) et des observations (données) avant et après traitement. Les deux groupes doivent être le plus homogène possible.

Dans cette étude, on assimile la survenue du cancer à une intervention (traitement) ; le groupe traité, aux individus avec une ALD30 sur la période étudiée (2005-2008) ; le groupe témoin, les individus sans ALD30 sur la période étudiée.

La base de données permet de recueillir pour chaque individu les données avant la survenue du cancer.

ii) que l'intervention (ici le cancer) soit attribuée (surviene) de façon aléatoire et exogène

Dans cette étude, la survenue du cancer peut être non exogène : certaines professions causant des cancers. L'endogénéité est traitée par les auteurs, en introduisant comme variable d'appariement (plus bas) la situation professionnelle, retardée (observée l'année précédant la survenue du cancer), pour contrôler pour les variables inobservables susceptibles d'influencer les parcours professionnels et de maladie.

L'effet du cancer est mesuré sur les parcours professionnels, caractérisés par l'emploi, le chômage, la maladie. Pour chaque type de parcours, trois indicateurs sont utilisés:

- Variable indicatrice = 1 si la personne a validé (au moins ?) un trimestre au régime général en emploi / chômage / maladie,

- Ratio : nombre de trimestres validés en emploi, chômage et maladie rapporté au nombre total de trimestres validés
- Nombre de trimestres validés

Si l'intervention n'est pas attribuée de façon aléatoire, l'appariement permet d'assurer l'homogénéité des deux groupes. Les variables d'appariement sont :

- 1) Date première inscription en ALD30, âge, genre et classe de salaire en début de fichier
- 2) Les 4 ci-dessus + profil d'emploi l'année avant la survenue du cancer (voir au-dessus)

- **Qualité de l'appariement excellente** : les taux d'appariement sont proches de 100% sans la variable emploi et de 98% avec l'emploi (tableau 3)
- **Caractéristiques en termes de statut sur le marché du travail des individus avec et sans ALD** (tableau 2). Si la proportion des personnes ayant au moins un trimestre d'emploi est équivalente dans les deux groupes, ce n'est pas le cas pour un trimestre en maladie : 8,7% traités contre 4,7% témoins ou le chômage : 12% traités, 15% témoins.
- **Effet du cancer sur les parcours professionnels** (emploi, chômage, maladie), appariement : genre, année de naissance, classe de salaire (tableau 4) à 1, 2, .. 5 ans

Question : la date de survenue du cancer n'apparaît pas, or il est dit au début que c'est une variable d'appariement

Les résultats en fonction du type d'appariement sont similaires

- Impact de la maladie sur les arrêts maladie augmente avec la survenue du cancer, l'effet est plus fort pour les femmes que pour les hommes : 7% avant, 51% après de femmes valident au moins un trimestre d'arrêt maladie contre 7% avant, 41% après pour les hommes
- Impact de la maladie sur l'emploi réel et plus fort pour les femmes que pour les hommes : le % de personnes employées durant un trimestre varie de 85% avant, 74% après, pour les femmes, de 87% avant, 78% après pour les hommes
- Impact de la maladie sur le chômage : diminution du % des personnes au chômage (et du nombre de trimestres) : passe de 18% à 13% pour les femmes, de 17% à 15% pour les hommes → substitution du chômage à arrêt maladie.

Les résultats en fonction du type d'appariement sont également similaires

- Impact de la maladie sur la probabilité d'être en arrêt maladie diminue régulièrement pendant les trois années après la survenue du cancer pour se stabiliser (0,44 1^{ère} année, 0,24, 0,003 en 5^{ième} année) de même que le nombre de trimestres en maladie. Pour les femmes par exemple, l'effet du cancer diminue de moitié dès la seconde année
- Impact de la maladie sur l'emploi : l'effet négatif est constant durant les cinq années tant pour les femmes que pour les hommes
- Au contraire, l'impact de la maladie sur le chômage disparaît après la première année, l'impact du cancer est donc très réduit sur cette variable.

Comparaison hommes-femmes indiquent que les hommes reprennent plus vite le travail que les femmes

Cancer : ne parle-t-on pas plutôt de cancers (voir le tableau 1)

- Présentation de la base de données (p5) n'est pas très claire pour un lecteur non spécialiste de ces bases de données.

On a parfois (au moins à la première lecture) que vous utilisez plusieurs bases de données, du coup on se perd - vous parlez de base Hygie, puis plus loin cette base de données (pour les ALD30), puis, bien que cette base et enfin en outre, nous disposons d'une base de données

Il suffit d'être plus précis, par exemple : « la base Hygie identifie les ADL et notamment l'ADL30. Elle comporte un certain nombre d'avantages, mais aussi des limites etc... ». L'avantage est sa couverture importante etc.. Sa limite est liée au fait que tous les cancers n'y sont pas répertoriés etc..

- **Présentation de la méthode de la double différence (p6)**

Il me semble qu'elle pourrait être présentée de façon beaucoup plus simple. La présentation que vous faites est compliquée ou du moins, et à mon avis, pas vraiment justifiée dans ce type d'article dans la mesure où cette méthode est bien connue et que par ailleurs, ce n'est pas l'objet de votre article que de présenter en détail les aspects techniques.

Il suffit de dire que cette méthode met en place deux groupes, un groupe qui a subi un choc de santé (équivalent d'un cas, dans la méthode d'analyse d'impact) et un groupe qui n'a pas subi ce choc (groupe témoin). Que pour estimer l'impact de ce choc, il faut observer le phénomène étudié avant et après la survenue du choc sur un groupe cas et sur un groupe témoin. Le groupe témoin devant être le plus proche possible du groupe traité.

C'est la comparaison des résultats économiques étudiés avant/après le traitement (ici, le choc de santé) dans deux groupes (traité, ici les malades, non traités, les sains) qui permet d'attribuer l'effet causal du choc de santé sur les variables expliquées (emploi, trajectoire professionnelle..).

Voir, Gertler et al, Sebastian Martinez, Patrick Premand, Laura B. Rawlings, Christel M. J. Vermeersch, L'Evaluation d'impact en pratique, Banque Mondiale, 2011, BIRD, BM.

Par contre, le modèle (équation) avec les différents paramètres qui seront estimés n'est pas présenté et cela manque, notamment pour comprendre ou mieux suivre les résultats des tableaux 4 et 5. Titre de ces tableaux, « La valeur moyenne de la variable - ajoutez parcours professionnel - pour les traités avant le cancer »

- Variables : leur présentation n'est pas très claire

Il faut relire plusieurs fois pour comprendre. Prenons le ratio : quand un individu ne valide pas un trimestre, tous les ratios sont égaux à 0 ? Est-ce plutôt quand un individu ne valide aucun trimestre, tous

Classe de salaire en début de fichier : que voulez-vous dire par début de fichier ??

- Résultats

Tableau 1, E11 = diabète, ce n'est pas un cancer, à enlever ?

Tableau 2 (p8) : vous ne présentez ni ne commentez la différence de % entre traités et témoins pour le chômage. Or, le fait que les témoins soient plus au chômage que les cas n'est pas a priori intuitive : vous pourriez évoquer déjà ici la substitution possible entre maladie et chômage

- Vous assimilez parcours professionnel et profil de carrière (ou évoquez l'un ou l'autre alternativement) à chômage, emploi, maladie. C'est un peu troublant au début, car on s'attend à autre chose lorsqu'on évoque généralement un profil de carrière
- Vous présentez des résultats uniquement en différenciant selon le genre. Au début, vous parliez de différencier aussi selon l'âge
- De même, vous évoquez au début des personnes en retraite, on n'en parle plus ensuite

Sur les groupes : vous comparez les parcours professionnels entre individus avec ALD30 et tous les autres, considérés comme témoins et que vous considérez comme en bonne santé. Or tous les témoins ne sont pas en bonne santé, certains ayant une ALD non cancéreuse, mais qui peut être aussi invalidante que l'ALD30.

N'est-ce pas un problème ?

- Biais de sous-estimation si l'ALD30 est la plus invalidante
- Biais car le groupe est hétérogène (individus peu malades et individus fortement malades sont ensemble) ? Vous n'en tenez pas compte dans l'appariement : les variables sont âge, sexe, emploi en t-1.
- N'aurait-il pas fallu apparier sur l'ALD (autre que ALD30) ?

- Ne serait-il pas intéressant de comparer l'impact du cancer et l'impact d'une autre ADL invalidante ? Ce qui implique de travailler sur un sous-échantillon (enlever les salariés sans ALD)
- Le tableau 2 montre que le % de personnes ayant validé un trimestre au chômage était plus faible avant la survenue de la maladie pour les personnes avec cancer (13%) que pour les non cancéreux (16%). Cette différence n'est pas prise en compte : ne peut-elle biaiser l'impact du cancer sur cette variable ?
 - Par ailleurs, vous n'appariez pas sur le type de cancers, n'est-ce pas gênant ? Car l'âge moyen de survenue des cancers ne diffère-t-il pas selon le type de cancer
 - L'impact en fonction de type de cancers serait intéressant, vous le prévoyez

- On ne sait rien des différences entre les individus par rapport au temps : certains seront jeunes, d'autres plus âgés. Si le cancer survient à 30 ans que se passe-t-il par rapport à un cancer qui survient à 50 ans
- Comment traitez-vous les personnes décédées ?