



Faire la sociologie des changements de pratiques avec une cohorte épidémiologique : l'apport de Constances

Marie Plessz

► To cite this version:

Marie Plessz. Faire la sociologie des changements de pratiques avec une cohorte épidémiologique : l'apport de Constances. Semaine Data-SHS, Progedo et Plateforme universitaire de données de Paris, Dec 2021, Paris, France. hal-03472394

HAL Id: hal-03472394

<https://hal.inrae.fr/hal-03472394>

Submitted on 9 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

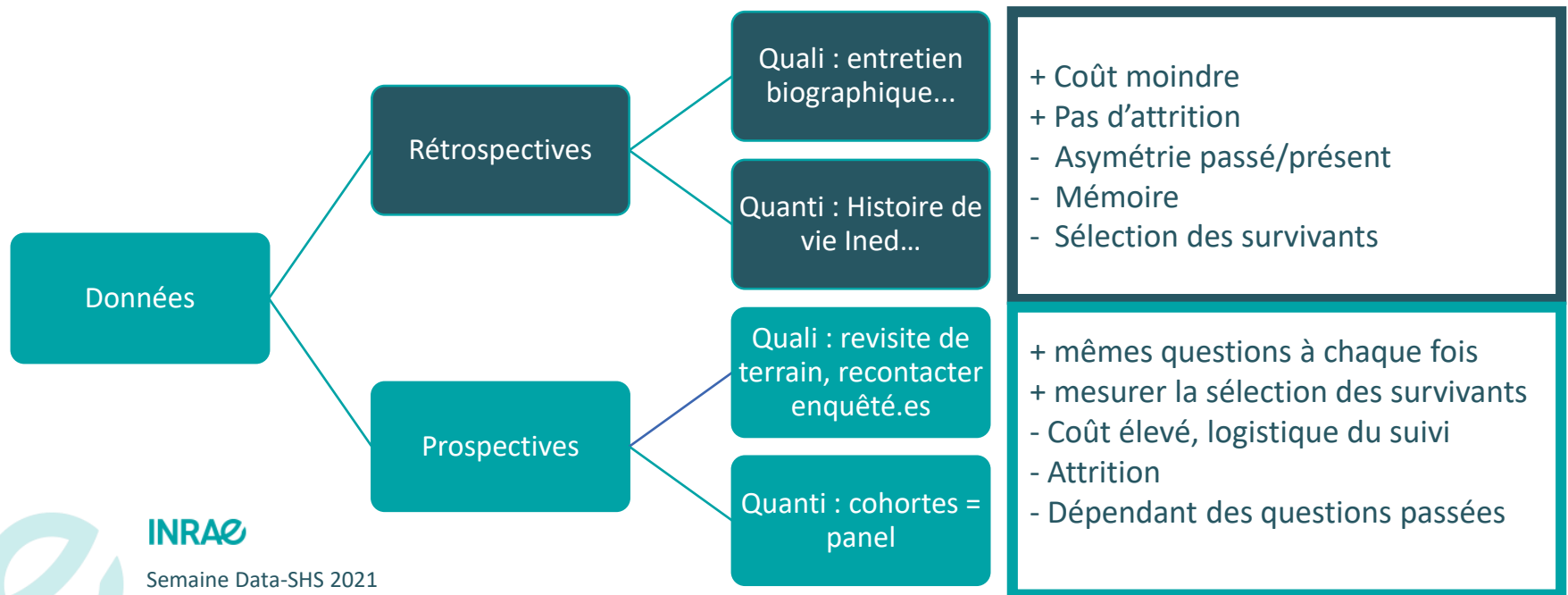
➤ Faire la sociologie des changements de pratiques avec une cohorte épidémiologique : l'apport de Constances

Marie Plessz, Centre Maurice Halbwachs

➤ Étudier les changements de pratiques

De la question de recherche aux données

- Comment changent les pratiques alimentaires?
 - Avancée en âge, différence entre cohortes de naissance, effet d'événements biographiques...
 - Aliments différents? Cuisiner plus, autrement? Manger seul·e? Chez soi?
 - Question supplémentaires : différences selon genre et position sociale



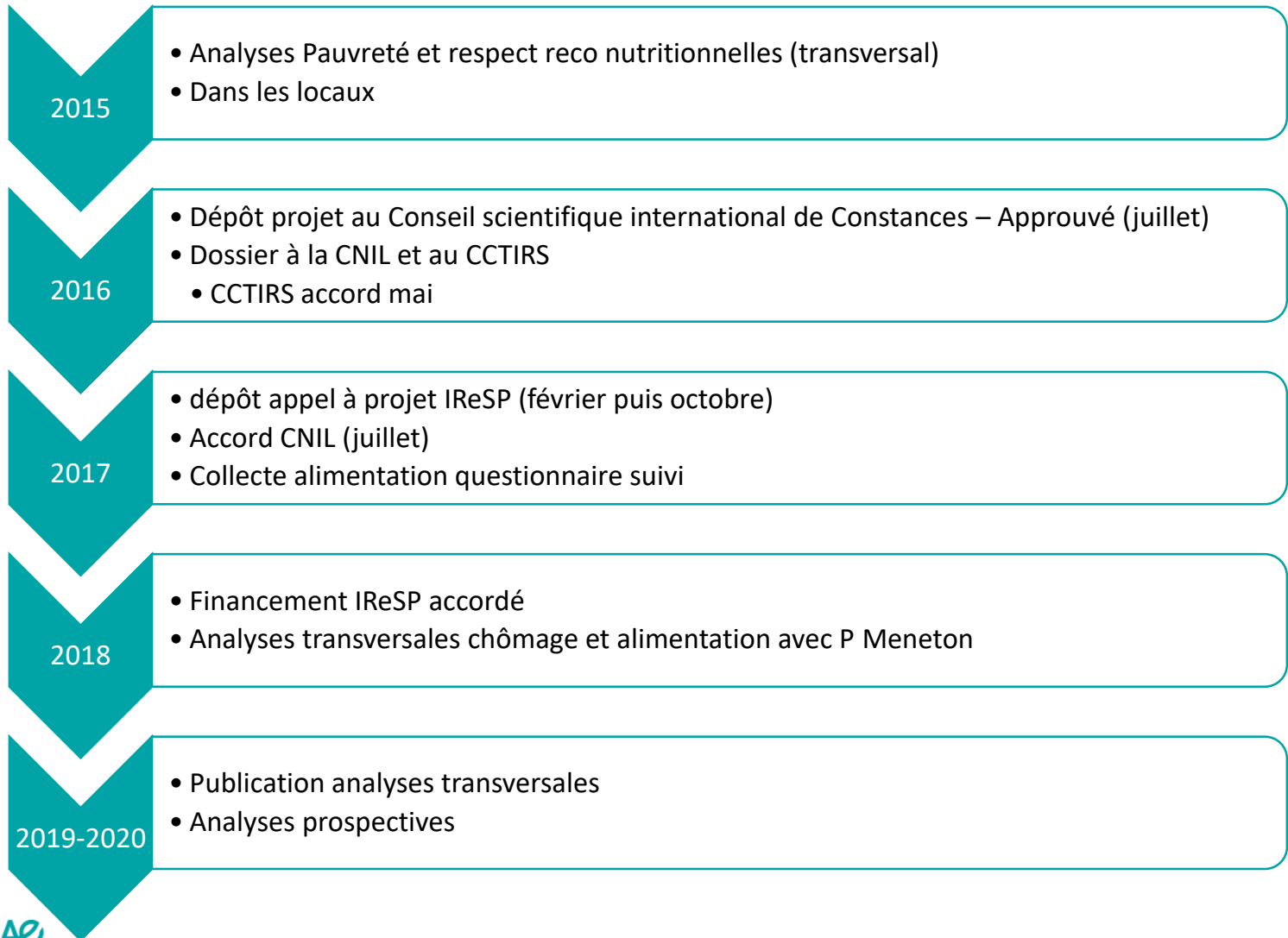
➤ Collaborer avec l'équipe, dans l'équipe

Séjour de recherche de 18 mois au sein de l'UMS (2013-2015)

- Contexte
 - Convention de mise à disposition (INRAE me paie pour travailler à l'INSERM)
 - Projet de recherche : consommation de légumes et vieillissement dans la cohorte GAZEL (appuyé sur mon travail en transversal déjà publié, *Sociology* 2015)
- Avantages
 - Possibilité d'interagir au quotidien
 - Excellente compréhension du processus de production des données
 - conseils des data managers, statisticiennes, doctorant-es, chercheurs, PU-PH
 - Données sécurisées sur les serveurs de l'UMS
- Limites
 - Capacités d'accueil
 - Travailler à 80 % sur Gazel ou Constances pendant plusieurs mois
 - INRAE ne fait plus de conventions de mise à disposition

➤ Un projet scientifique avec les données Constances

CALICO (chômage et alimentation dans Constances)



➤ Des données superbes mais complexes

- Gros travail de nettoyage des données par l'équipe de constances
 - Cohérence entre filtres et questions filtrées
 - Cohérence entre versions papier et web
 - Documenté dans application AQUEDUC
- Métadonnées
 - Tableau excel listant les variables de chaque table
 - Fichier SAS pour ajouter des formats à chaque variable, en français ou en anglais
- Variables « Q°r_rep » pour chaque questionnaire
 - Pour identifier les individus qui ont rendu chaque questionnaire
- Plusieurs tables de données
 - Par source (questionnaire, calendrier etc)
 - Par structure (nb de lignes/individu)
- Interpréter les cases vides
 - Question filtrée
 - Questionnaire non rendu/pas encore arrivé/intégré dans la base
 - Non-réponse
- Complexité des données de panel
- Assurer la sécurité des données
 - toutes les opérations sont réalisées en lignes de commandes dans des programmes
 - Ne pas multiplier les sauvegardes des données

➤ Quelques trucs et astuces

Organisation du travail et des programmes

- Planifier

- Lister les variables dont on pense avoir besoin,
- repérer dans quels fichiers de données elles sont
- identifier la structure du fichier adéquate pour les analyses (une ligne par individu? 1 ligne/individu/année? 1 ligne/individu/année/aliment?)
- Créer le dossier du projet avec les sous-dossiers de son choix

- Plan de travail possible

1. Réaliser les recodages sur chaque fichier de données : MDV, POP etc
 - Vérifier les recodages en croisant avant/après
2. Conserver seulement les variables utiles + identifiants individuels et Q^or_rep
3. Apparier tous les fichiers (merge, join)
4. Sélectionner la population d'étude.
 - Vérifier la distribution des variables avant/après suppression des cas incomplets.
 - Réfléchir aux biais de non-réponse
5. Éventuellement quelques recodages supplémentaires
6. Analyses

- Tester les manipulations complexes sur quelques individus

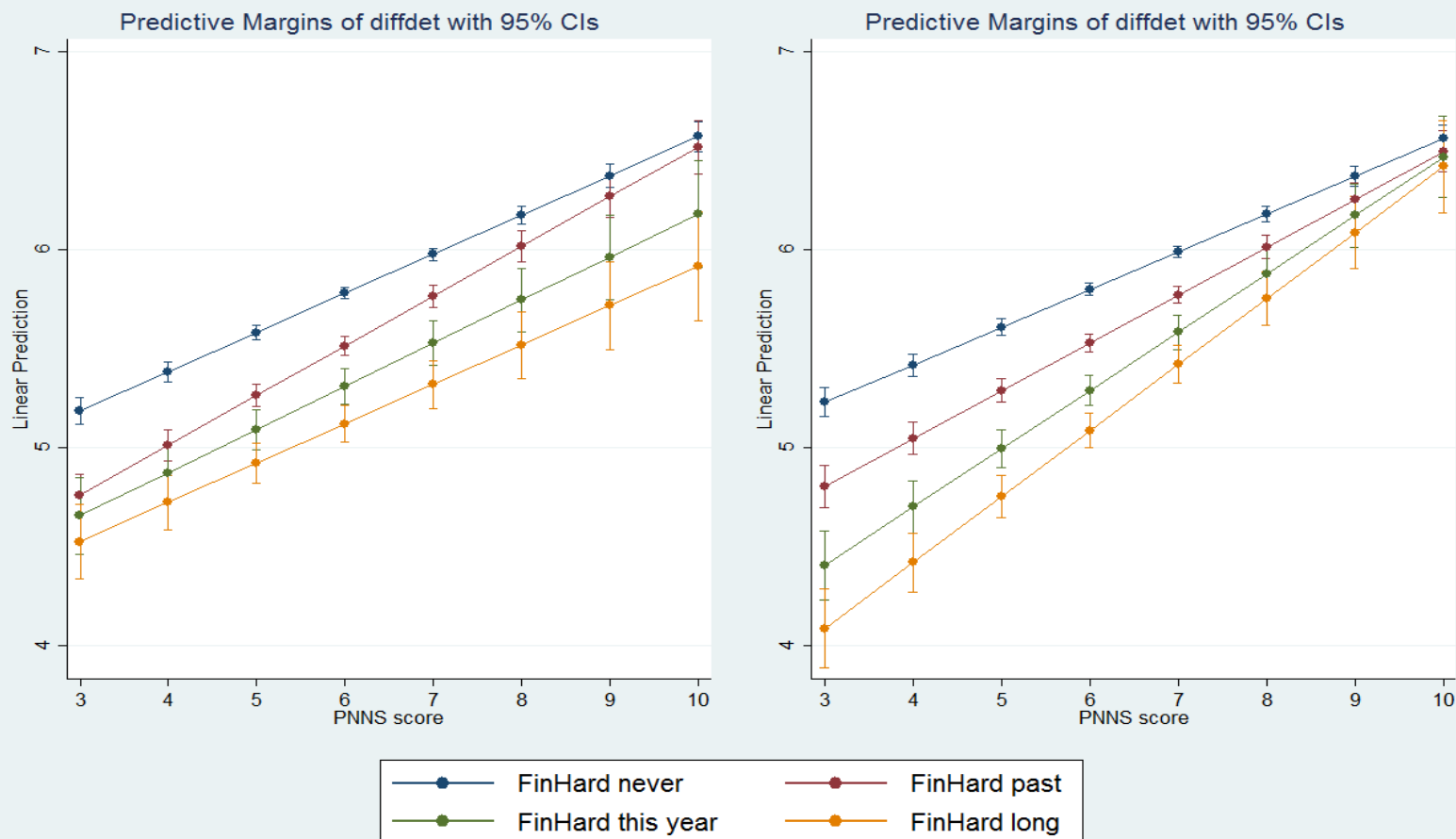
- Appariement de tables (merge), restructuration (reshape, pivot, melt/dcast etc)

➤ Quelques résultats

Les personnes en difficulté financière ne surestiment pas la qualité de leur alimentation (Plessz et al, *Appetite*, 2019)

Score PNNS et équilibre alimentaire perçu selon les difficultés financières

Prédiction linéaires pour les hommes (gauche) et les femmes (droite)



Régression linéaire sur données imputées. Contrôles: en couple, régime prescrit, âge, diplôme. Constances, MP, 14 Apr 2017.

➤ Quelques résultats

Par rapport aux personnes en emploi, les chômeurs ont nettement plus de pratiques à risque, mais pas plus que si les pratiques étaient indépendantes les unes des autres (Plessz et al, *PLoSone*, 2020)

Table 2: Relative risk ratios of co-occurring risk factors according to experience of unemployment in Const (reference: 0 risk).

No. of risk factors	Experience of unemployment	RRR	M1 95% CI	p	RRR	M2 95% CI	p
1	Never	1			1		
	Past only	1.07	0.99 - 1.15	0.077	0.98	0.92 - 1.06	0.67
	Now only	1.17	1.03 - 1.34	0.018	1.07	0.94 - 1.22	0.33
	Past & now	1.11	0.92 - 1.34	0.280	1.00	0.83 - 1.22	0.98
	<i>p for trend</i>			0.172			0.76
2	Never	1			1		
	Past only	1.3	1.21 - 1.40	<0.0001	1.12	1.04 - 1.21	0.003
	Now only	1.65	1.45 - 1.89	<0.0001	1.34	1.17 - 1.54	0.00002
	Past & now	1.82	1.50 - 2.20	<0.0001	1.46	1.20 - 1.77	0.0001
	<i>p for trend</i>			<0.0001			<0.0001
3	Never	1			1		
	Past only	1.62	1.48 - 1.78	<0.0001	1.29	1.17 - 1.42	<0.0001
	Now only	2.51	2.16 - 2.93	<0.0001	1.67	1.43 - 1.95	<0.0001
	Past & now	2.47	1.98 - 3.08	<0.0001	1.65	1.31 - 2.07	<0.0001
	<i>p for trend</i>			<0.0001			<0.0001
4	Never	1			1		
	Past only	1.86	1.53 - 2.25	<0.0001	1.40	1.15 - 1.69	0.001
	Now only	2.95	2.26 - 3.86	<0.0001	1.71	1.30 - 2.25	0.0001
	Past & now	3.74	2.59 - 5.40	<0.0001	2.19	1.51 - 3.19	<0.0001
	<i>p for trend</i>			<0.0001			<0.0001

Multinomial regression. RRR: relative risk ratio. The base level of the outcome is: exposed to zero risk.

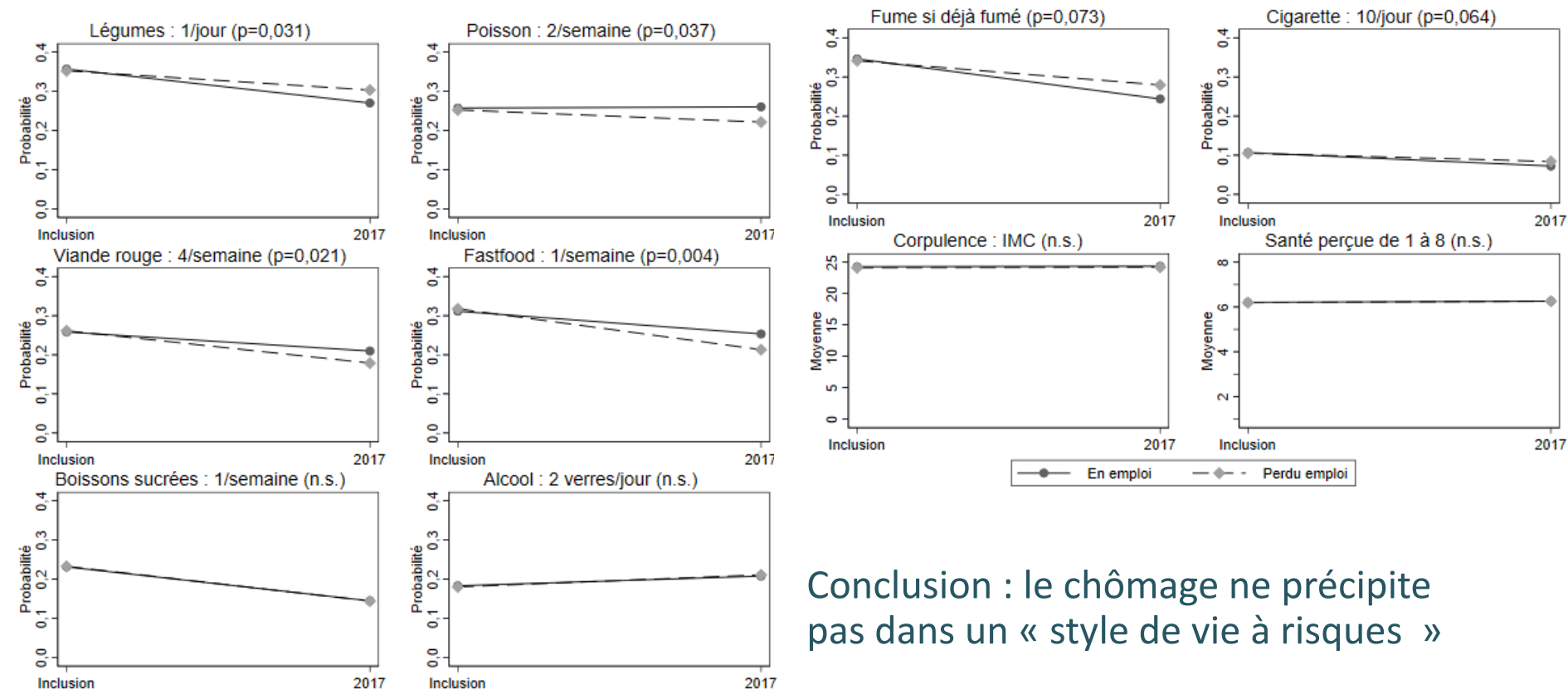
M1: adjusted on sex and age

M2: M1 + self-rated health, education, partnership status, presence of children, region



➤ Quelques résultats

La consommation d'aliments, de boissons et de tabac de tous les enquêtés évolue entre l'inclusion et le suivi de 2017. Les différences liées à la perte d'emploi sont minimales et non systématiques



Conclusion : le chômage ne précipite pas dans un « style de vie à risques »

> Conclusion

Mon dialogue entre SHS et Santé publique/épidémiologie

- Facilité par
 - Bonne maîtrise d'un logiciel de statistiques
 - Une école d'été INSERM sur les données longitudinales
 - La culture « épidémiologie sociale » très forte dans l'équipe Constances
- Les points de discussion
 - Questions « validées » vs questions testées
 - Exposition/contrôle/outcome vs décrire des trajectoires
 - Représentativité vs population cible
 - Qui est auteur?
 - ➔ Occasions de mieux comprendre comment fonctionnent nos disciplines
- Les points de rencontre
 - Données manquantes
 - Prendre en compte les conditions de remplissage du questionnaire (longueur, complexité)
 - Importance de la qualité des données et de l'échantillonnage

Plus de réflexions, de socio, de tableaux, programmes Stata et R dans mon HDR
<https://hal.inrae.fr/tel-03436087> et sur gitlab <https://forgemia.inra.fr/marie.plessz>