

Données DSN et provisionnement de l'arrêt de travail : Une application pratique

Hicham BOUDDOUR

Responsable Pôle Actuariat
et Souscription

IPECA - MSAé

Emmanuel GINESTE

Directeur de mission
Actélior



- Institution de prévoyance, IPECA a fêté ses 75 ans le 7 octobre dernier
- Couvre 6 salariés sur 10 de l'aéronautique et du spatial civils et militaires
- IPECA en chiffres :
 - ✓ 226 M€ de chiffre d'affaires
 - ✓ 271% de ratio de couverture du SCR
 - ✓ 418 entreprises adhérentes
 - ✓ 287 000 bénéficiaires, dont 36 000 retraités
 - ✓ 119 salariés



- Cabinet d'actuares conseils fondé en 1999, spécialisé en assurance de personnes
- 30 consultants et plus de 150 clients actifs

Plan

1. Introduction
2. Présentation du portefeuille
3. Données DSN
 - a. Périmètre et limitations
 - b. Traitements des données
 - c. Réconciliation avec les données assurantielles
4. Utilisation des données DSN pour le provisionnement
 - a. Méthode et hypothèses
 - b. Construction des triangles
 - c. Résultats
 - d. Pistes d'optimisation
5. Conclusion
6. Questions / Réponses

1. Introduction

Cet atelier a pour objectif de présenter les travaux menés par IPECA en 2022, en collaboration avec le cabinet d'actuares conseils Actélior, sur l'utilisation des données DSN. Ces travaux avaient deux objectifs :

- ✓ Editer des statistiques sur l'absentéisme de son portefeuille
- ✓ Elaborer une méthode alternative de provisionnement pour les provisions pour sinistres inconnus en incapacité

L'atelier permettra de présenter les travaux relatifs au deuxième objectif

1. Introduction

- La Déclaration Sociale Nominative (DSN) est le moyen de transmission des données employeurs pour déclarer et payer leurs cotisations aux organismes sociaux
- Elle est mensuelle et comprend des informations issues de la paie au sens large et les signalements d'évènement pour l'assurance maladie
- Elle a été mise en œuvre progressivement depuis 2017 et est généralisée depuis le 1er janvier 2019
- Elle est déclarée mensuellement, voire de façon infra-mensuelle pour certaines évènements tels qu'un arrêt de travail
- En pratique, elle est constituée d'une vingtaine de tables, chacune contenant des informations spécifiques sur la situation des salariés

2. Présentation du portefeuille

- Pour une plus grande homogénéité des données, les travaux ont porté sur le périmètre AIRBUS uniquement
- Nous avons considéré les arrêts survenus sur la période 2019-2021
- Le portefeuille compte 54.000 têtes, avec un peu plus de 1.000 arrêts indemnisés par an
- La population de ce portefeuille est relativement stable sur la période considérée, tant en effectifs qu'en caractéristiques (sexe, âges, catégories ...)
- L'indemnisation d'IPECA intervient en relais des obligations de maintien de salaire de l'employeur (CCN + accord collectif). Le maintien CCN à 100% est de 3 mois à partir d'un an d'ancienneté est atteint 6 mois à partir de 15 ans d'ancienneté. L'accord collectif peut prévoir un maintien au-delà des 6 mois. Après le maintien à 100%, un maintien à 50% est également prévu pour une durée dépendant aussi de l'ancienneté
- Les accidents du travail et les maladies professionnelles ne sont pas pris en charge par l'IPECA

3.a Données DSN - Périmètre et limitations

- Les données DSN sont considérées comme fiables à compter de janvier 2019. Nous disposons des données 2018, elles ne sont cependant pas utilisées, sauf pour les retraitements (cf. ci après)
- La période d'étude est donc relativement courte

3.a Données DSN - Périmètre et limitations

- Les données DSN comprennent :
 - La date de naissance
 - Le sexe
 - Le collège
- Elles sont limitées sur la cause :
 - Maladie
 - Temps partiel thérapeutique
 - Maladie professionnelle
 - Accident du travail ou du trajet
 - Autres causes (paternité, maternité ...)

3.b Données DSN - Traitement des données

- Les données DSN sont plutôt propres malgré un déficit sur l'alimentation de l'item « sexe »

Nombre de périodes DSN	Pourcentage de périodes DSN pour lesquelles le sexe n'est pas renseigné	Pourcentage de périodes DSN pour lesquelles la date de naissance n'est pas renseignée	Pourcentage de salariés pour lesquels le collège n'est pas renseigné
228 194	11.2%	0.5%	0.6%

- Les données de gestion d'IPECA permettent de réduire très sensiblement le nombre de valeurs manquantes
- Moins de 0,1% des lignes sont finalement exclues

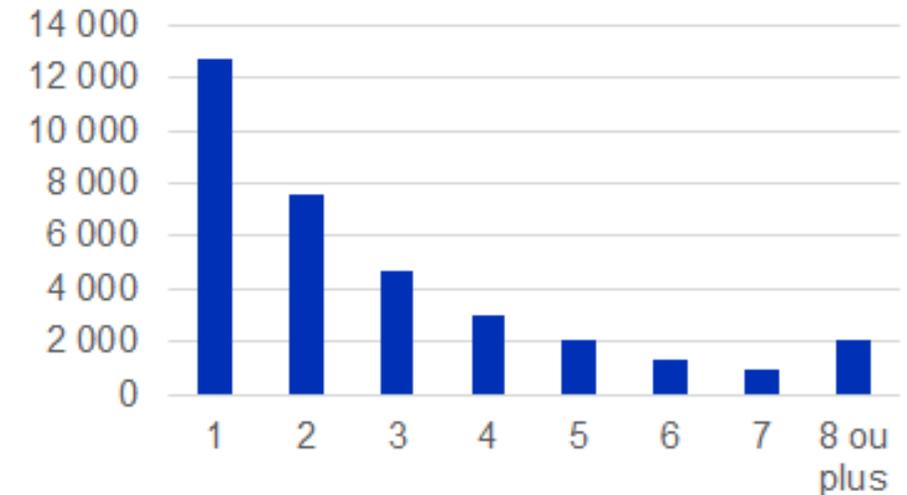
3.b Données DSN - Traitement des données

- La table des arrêts DSN utilisée se présente sous la forme d'une vision cumulée de toutes les lignes DSN reçues à la date d'extraction. Pour un même arrêt, nous disposons donc des informations des dates de fin de période à différentes dates. Seule la ligne correspondant à l'information la plus récente est conservée
- Pour notre étude nous n'avons conservé que les causes prises en charge par l'IPECA : maladie et temps-partiels thérapeutique

3.c Données DSN - Réconciliation avec les données assurantielles

- La réconciliation s'avère problématique du fait de logiques différentes entre la base DSN et la base IPECA :
 - En cas de reprise d'activité entre deux périodes d'arrêts, la DSN contient deux lignes. L'absence d'information précise sur la cause ne permet pas d'identifier s'il s'agit d'une rechute ou d'une nouvelle cause
 - Pour le portefeuille IPECA, pour la même cause, si la reprise est inférieure à 12 mois on considère qu'il s'agit du même sinistre
 - Les tentatives de regroupement des sinistres DSN afin de valider le nombre de sinistres par rapport à celui des bases IPECA n'ont pas abouti. Cela s'explique pour partie par le grand nombre de sinistrés ayant de multiples périodes d'indemnisation dans la base DSN.

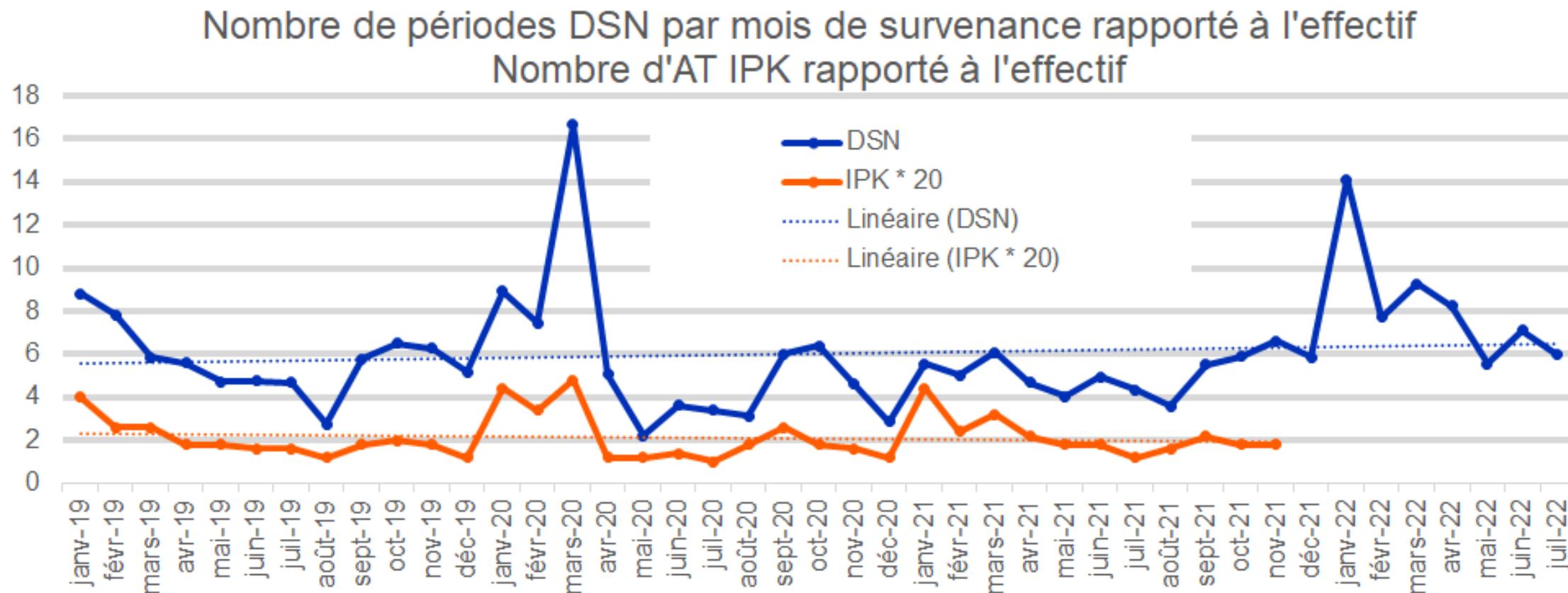
Effectifs des sinistrés selon le nombre de lignes d'arrêts DSN sur la période 2019-2021



Le maximum est de 33 périodes d'arrêts DSN pour un même assuré, entre 2019 et 2021 ...

3.c Données DSN - Réconciliation avec les données assurantielles

- Le lien entre les données DSN et la sinistralité IPECA est cependant établi :



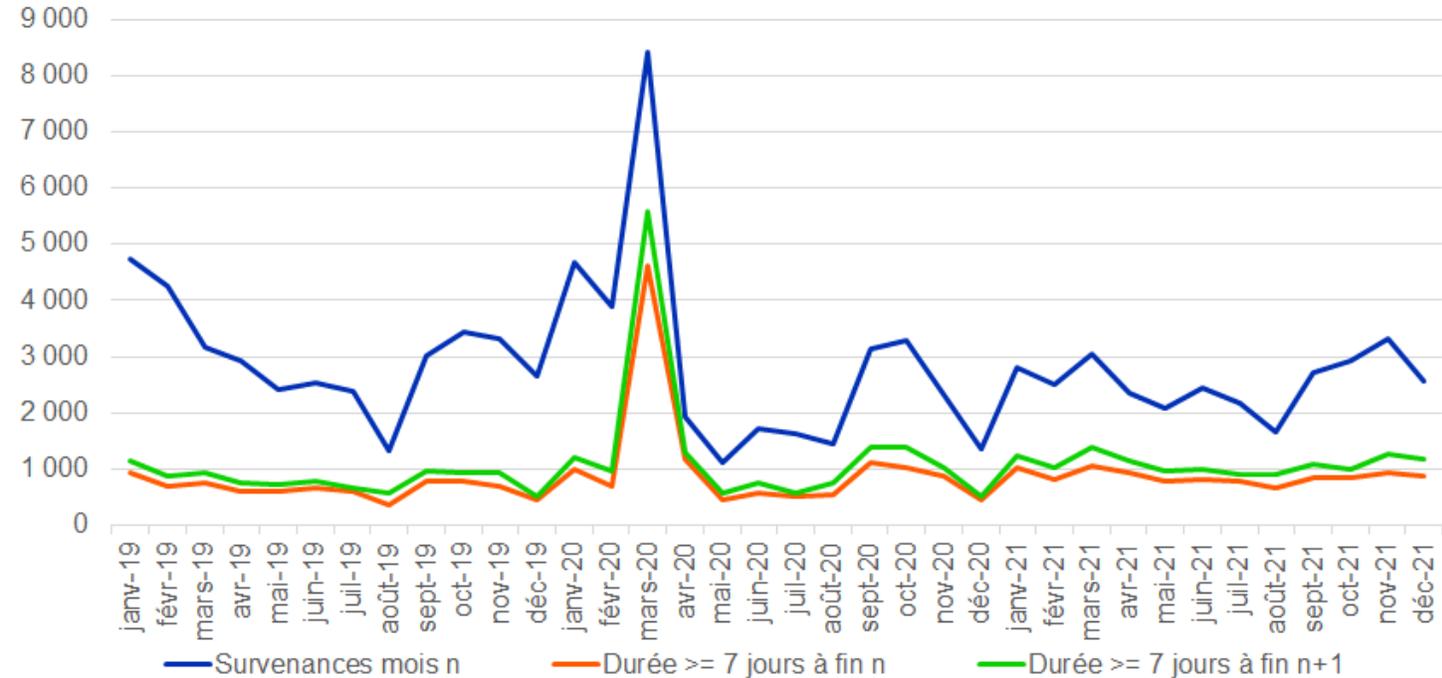
4.a Provisionnement – Méthode et hypothèse

- L'objectif poursuivi est d'estimer les IBNR, a minima pour les derniers mois de l'année de la clôture, pour le risque incapacité. La maille choisie est mensuelle

- Plusieurs tentatives ont été effectuées pour trouver une corrélation entre le nombre de sinistres indemnisés par IPECA et le nombre de lignes DSN :

✓ Lignes DSN regroupées

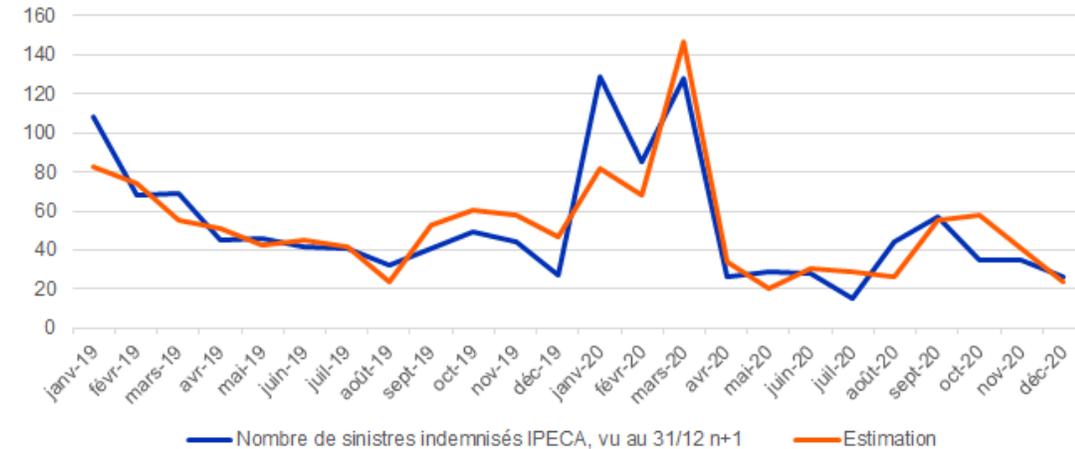
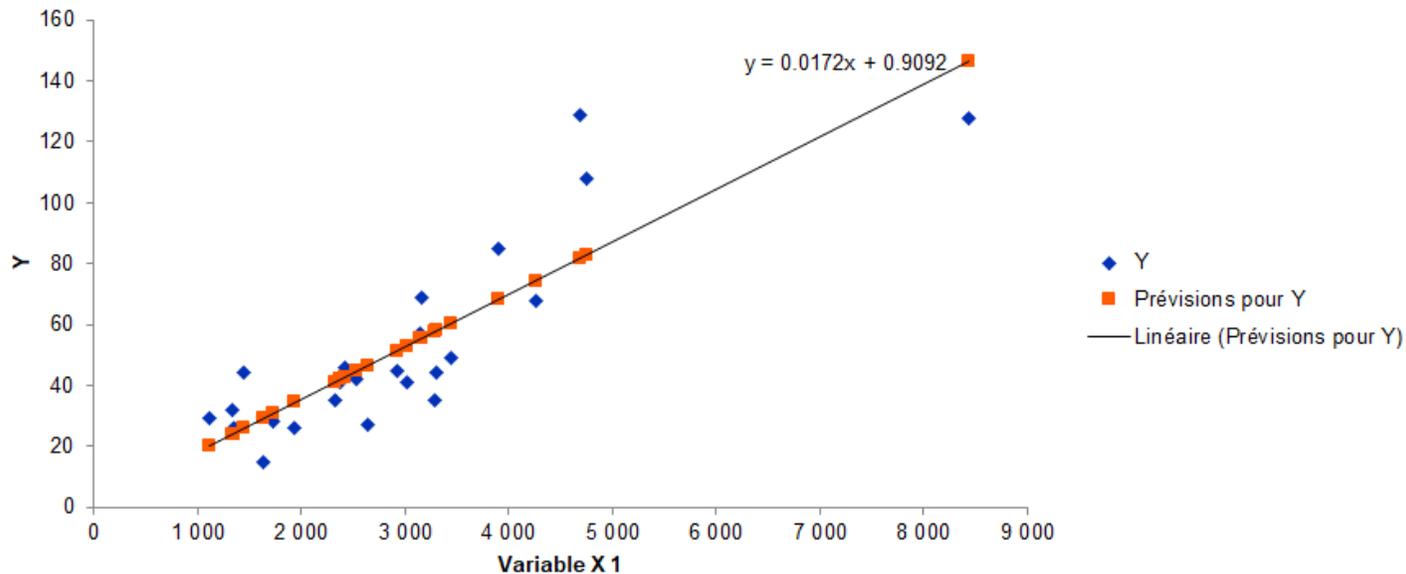
✓ Lignes DSN non regroupées, vu à fin du mois de survenance, au mois suivant, pour les arrêts ayant une durée de plus de 7, 14 ou 30 jours



4.a Provisionnement – Méthode et hypothèse

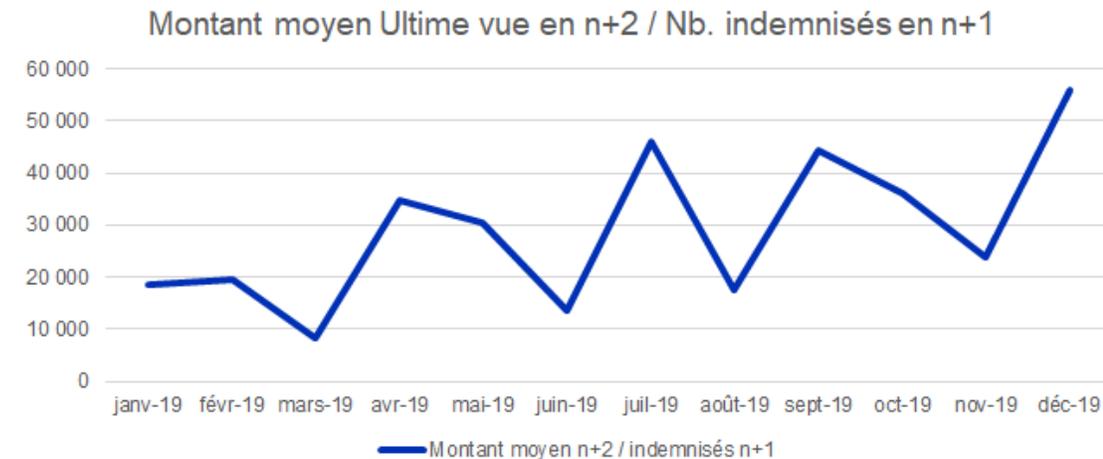
- Les résultats les plus probants sont finalement obtenus avec les données demandant le moins de retraitement : le nombre de lignes DSN survenues le mois n, vu à la fin du mois n

Variable X 1 Courbe de régression



4.a Provisionnement – Méthode et hypothèse

- Passage du nombre de sinistres à la provision :
 - Les montants moyens par mois sont volatils
 - L'estimation de cette valeur, vue au 31/12/N pose clairement problème
 - Cette série est de plus perturbée pour l'exercice 2020



4.c Provisionnement – Résultats

- Peu de recul pour valider le bien fondé de la méthode
- Au regard de l'exercice 2019, la méthode aurait conduit à sur-provisionner : +11%
- La méthode reste donc à perfectionner

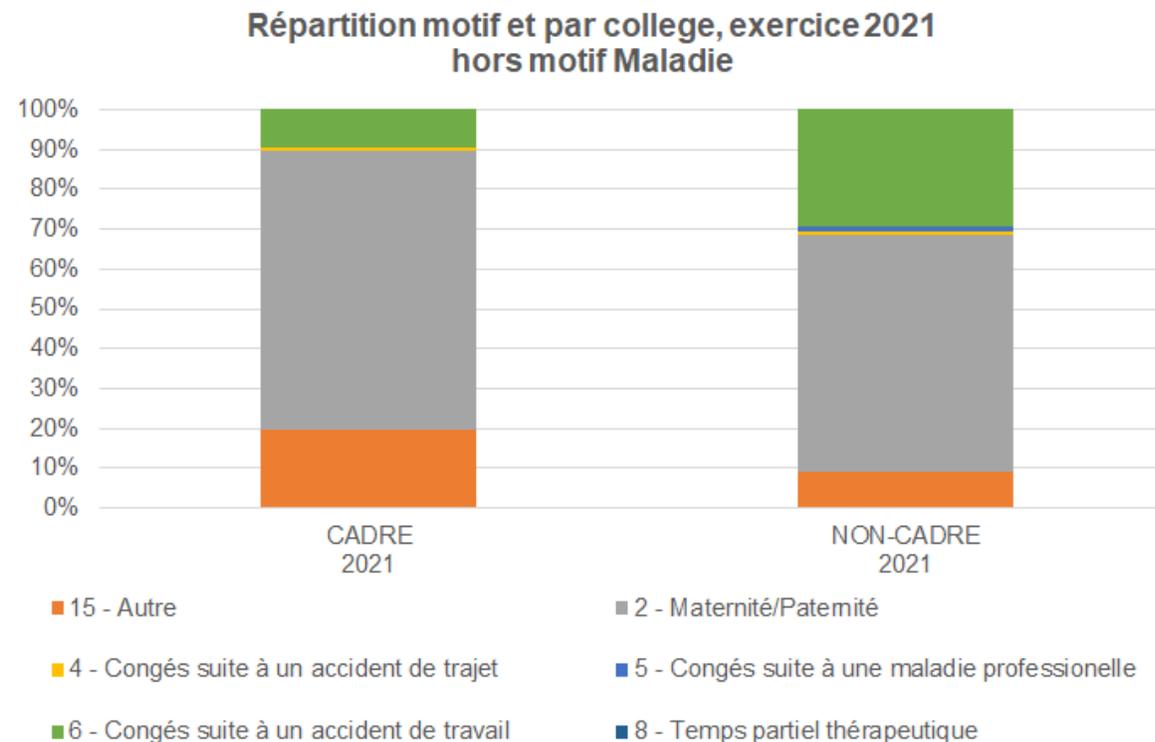
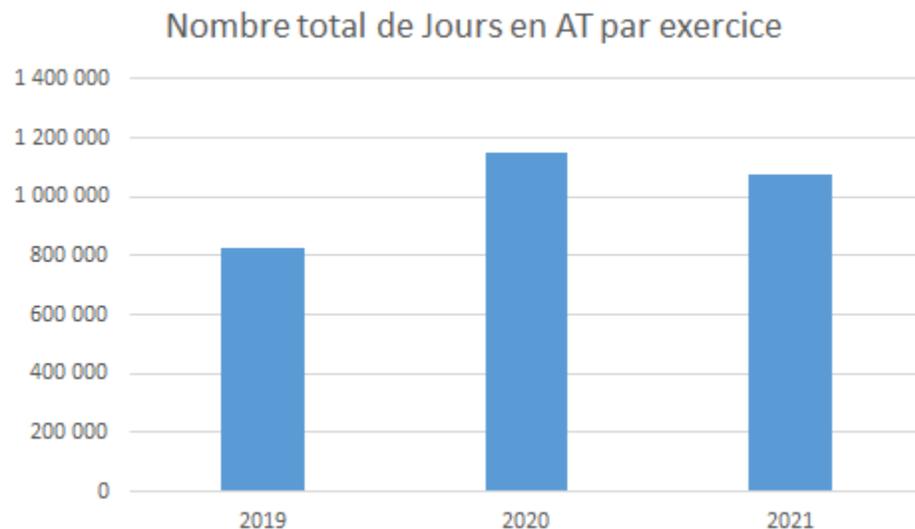
	2019
Montant réglé + PM au 31/12/N	22 676 k€
Estimation Ultime montant moyen	20 715 k€
Estimation Ultime montant moyen 6-6	20 033 k€
Dernière situation connue	18 040 k€

4.d Pistes d'optimisation

- La première approche a été plutôt basique avec la prise en compte de l'ensemble du portefeuille, sans segmentation
- Une segmentation permettrait probablement d'obtenir des résultats plus précis
 - ✓ Bien que les segmentations possibles soient limitées par les données disponibles dans la DSN, celles qui paraissent les plus naturelles sont accessibles :
 - Par collègue
 - Par tranche d'âge
 - (Par sexe)
 - ✓ Il convient cependant de veiller à ce que la segmentation ne soit pas trop fine au regard de la taille du portefeuille et du nombre d'observations disponibles
- La cadence des survenances met en évidence une périodicité infra-annuelle marquée, avec notamment un pic en janvier. Cette périodicité n'a pas été exploitée à ce stade

5. Conclusion

- Les données DSN constituent une mine d'information, en particulier pour éditer des tableaux de bords de suivi de l'absentéisme et des éventuelles dérives



5. Conclusion

- Les données DSN constituent un indicateur avancé pertinent de la sinistralité, en particulier sur les contrats à franchises longues comme pour le portefeuille de l'IPECA
- Les données DSN sont cependant délicates à manipuler et une bonne compréhension de leur structure est nécessaire
- Pour notre étude, la période limitée pour les données disponibles et le contexte spécifique pour une partie significative de cette période (crise sanitaire) nourrissent des craintes sur la robustesse de la méthode employée
- L'étude a également permis de confirmer l'allongement de la durée moyenne des arrêts de travail
- L'information limitée sur la cause de l'arrêt ne permet pas des études plus fines sur certaines cause d'arrêts en croissance, telles que les affections psychologiques

Des questions ?

