

# **Rapport de Production**

## **Années de décès : 2018-2019**

### **Données définitives**

Zina Hebbache, Aude Robert, François Clanché (\*),  
Elise Coudin, Diane Martin  
CépiDc-Inserm, (\*) DREES

Septembre 2023

**Document de travail du CépiDc N°3**

*Ces documents de travail ne reflètent pas la position de l'Inserm et n'engagent que leurs auteurs.*

## Table des matières

<b>1</b>	<b>INTRODUCTION.....</b>	<b>3</b>
<b>2</b>	<b>COLLECTE .....</b>	<b>4</b>
2.1	EXHAUSTIVITE .....	4
2.2	ORIGINE DES CERTIFICATS .....	5
2.3	VERSION DE CERTIFICATS .....	5
2.4	VOLETS MEDICAUX COMPLEMENTAIRES (VMC) .....	5
2.5	TYPE DE CERTIFICAT .....	6
<b>3</b>	<b>CODAGE .....</b>	<b>6</b>
3.1	CONTEXTE ET STRATEGIE .....	6
3.1.1	<i>Modes de codage</i> .....	6
3.1.2	<i>Échantillons de codage manuel</i> .....	7
3.1.3	<i>Processus de vérification</i> .....	9
3.2	EVOLUTIONS DE CODAGE .....	13
3.2.1	<i>Nouveautés relatives aux recommandations OMS</i> .....	13
3.2.2	<i>Précisions dans l'application des règles de l'OMS</i> .....	14
3.2.3	<i>Mises à jour du dictionnaire des expressions nosologiques utilisé par Iris Muse</i> .....	14
<b>4</b>	<b>EVOLUTION DE LA BASE DE DONNEES.....</b>	<b>15</b>
4.1	NOUVELLE VERSION DE CERTIFICATS DE DECES .....	15
4.2	NOUVEAU MODE DE CODAGE .....	15
<b>5</b>	<b>ANNEXE .....</b>	<b>16</b>
5.1	VERSION 2017 DU VOLET MEDICAL DU CERTIFICAT DE DECES GENERAL (28 JOURS ET PLUS).....	16
5.2	VERSION 2017 DU VOLET MEDICAL DU CERTIFICAT DE DECES NEONATAL (MOINS DE 28 JOURS) .....	17
5.3	CERTIFICATS DES DECES ENTRE 28 JOURS ET 15 ANS IDENTIFIES DANS LES DECES SENSIBLES HORS MORTS VIOLENTES ....	18
5.4	DICIONNAIRE DE VARIABLES DANS LE SNDS.....	19
5.5	LISTE DES TABLEAUX .....	26

## 1 Introduction

Le processus de production des données du CépiDc est décrit en détail dans le document 'Statistiques sur les causes de décès de A à Z' accessible sur le site internet du CépiDc.

Ce rapport présente les spécificités du traitement des certificats de décès d'une année donnée et apporte des précisions sur le processus de production. Exceptionnellement ce rapport est commun aux années 2018 et 2019.

La base de données pour les décès 2018 et 2019 a été produite dans un contexte de rattrapage du retard de production et au moment où de nouvelles méthodes de codage automatique impliquant des algorithmes de deep learning ont été développées au CépiDc. Ainsi, le processus classique de codage a été adapté et les nouvelles méthodes de codage automatique ont été utilisées pour permettre de produire la base plus rapidement que les autres années. Ce rapport décrit comment le codage a été réalisé et permet d'avertir les utilisateurs sur les écarts entre la campagne de codage en rattrapage de 2018 et 2019 et une campagne traditionnelle de codage ne mobilisant que le système de règles et le codage manuel.

Ce rapport est à associer au document Codage des causes de décès de 2018 et 2019 en CIM10 - Approche combinant deep learning, système expert et codage manuel ciblé », document de travail du CépiDc n°2, septembre 2023 ([disponible sur le site internet du CépiDc](#)) qui présente en détail la combinaison des trois modes de codage lors de la campagne 2018 et 2019 et évalue son niveau de cohérence avec une campagne traditionnelle.

## 2 Collecte

### 2.1 Exhaustivité

Concernant les décès de l'année 2018, 610 136 décès ont été comptabilisés par l'Insee sur le territoire français. Le CépiDc a reçu 606 157 certificats et en a conservé 594664 volets médicaux (soit 98,1%). Pour 2,5% des décès, aucun volet médical n'a été reçu, et donc aucune cause n'est déclarée.

En 2019, 613 867 décès ont été comptabilisés sur le territoire français. Le CépiDc a reçu 608 458 volets médicaux et en a conservé 598707 (soit 98,4%). Pour 2,5% des décès, aucune cause n'a été déclarée.

En 2018, 122 des volets médicaux de décès de moins de 28 jours n'ont pas été reçus et 135 en 2019.

Tableau 1. Exhaustivité de la collecte des volets médicaux (VM) pour les décès 2018 et 2019

	<b>2018</b>	<b>2019</b>
Décès Insee	610136	613867
VM reçus	606157	608458
VM supprimés	-11493	-9751
VM non reçus	15472	15160
% de VM non reçus	2,54%	2,47%

Les volets médicaux supprimés correspondent majoritairement à des doublons (le décès a fait l'objet de plusieurs volets médicaux). Il y a aussi quelques cas où des certificats n'ont pas été appariés à un décès enregistré par l'Insee. Il peut s'agir de volets médicaux pour lesquels les données individuelles étaient incomplètes ou erronées (le certificateur a pu avoir fait un autre volet médical, c'est donc similaire à un doublon) ou encore il peut s'agir de volets médicaux tests (des tests sur l'application de certification électronique par exemple), mais ce dernier cas reste très anecdotique.

Les décès pour lesquels aucun volet médical n'a été reçu seront présents dans la base de données du CépiDc et leur cause initiale sera R99 (cause indéterminée). Tous les décès dont la cause initiale est R99 ne sont pas forcément des certificats non reçus au CépiDc. Il peut simplement s'agir de certificats dont la partie cause de décès n'a pas été complétée par le médecin ou insuffisamment pour identifier une cause initiale de décès informative. Pour repérer la différence de situation entre les certificats non reçus et ceux dont la cause était trop peu informative pour être codée autrement qu'en R99, il faut regarder s'il y a la mention « pas de certificat » dans les causes. En 2018, 56% des R99 étaient des certificats non reçus ; en 2019 seulement 48% des codes R99 correspondaient à un certificat non reçu. La forte augmentation des codes R99 parmi les certificats reçus en 2019 est due à une évolution dans les règles de codage de l'Organisation mondiale de la santé (OMS) : les codes I460 et I469 ne sont plus acceptés en cause initiale, car trop peu informatifs et sont remplacés par R99 à partir de 2019.

## 2.2 Origine des certificats

Pour l'année 2018, sur l'ensemble des volets médicaux reçus et conservés au CépiDc, 500 078 (84%) étaient au format papier, 94 586 (16%) étaient d'origine électronique.

Pour l'année 2019, sur l'ensemble des volets médicaux reçus et conservés au CépiDc, 484 198 (81%) étaient au format papier, 114 509(19%) étaient d'origine électronique.

Les causes rédigées par le médecin sur les certificats papiers sont saisies par un prestataire qui applique des règles de saisie précises permettant de faciliter leur codage automatique (corrections de fautes d'orthographe, suppression des articles...). A partir des décès 2018, ces règles de standardisation sont également appliquées sur les causes des certificats électroniques rejetés par le système expert de codage. Dans ce cas, le prestataire de saisie a appliqué les mêmes règles de standardisation que celles appliquées sur les formulaires papier.

## 2.3 Version de certificats

En 2016, lors de la parution du volume 2 de la classification internationale des maladies (CIM), qui est le volume comprenant toutes les nouvelles règles de codage, l'OMS a proposé un nouveau modèle de certificat. La France a mis en place ce certificat en 2018.

Ainsi, à partir du 1<sup>er</sup> janvier 2018, une nouvelle version du certificat ("certificat 2017", conforme à l'arrêté du 17 juillet 2017 <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000035388290>) a été mise en circulation. Celle-ci contient des informations complémentaires concernant notamment les circonstances apparentes de décès. Le médecin peut désormais renseigner s'il semble s'agir d'un suicide, un accident, une atteinte à la vie d'autrui, une mort naturelle (maladie) en cochant une case. De plus, dans la nouvelle version du certificat néonatal, le certificateur peut renseigner l'enchaînement causal ayant conduit au décès. Sur la version de 2017, il ne pouvait qu'indiquer la cause ayant directement provoqué le décès et d'autres causes indirectement impliquées dans le décès. Il y a également la notion de mort subite du nourrisson.

L'application de certification électronique a mis en place la nouvelle version de certificat le 02/01/2018 au matin. En papier, il persiste dans les faits une cohabitation des deux versions (1997/2017) pour des raisons pratiques et logistiques (les médecins utilisent les stocks de certificats ancien modèle encore en leur possession) pendant plusieurs années.

Sur l'ensemble des volets médicaux reçus et conservés au CépiDc pour l'année 2018, 305 809 (51,4%) correspondaient à la nouvelle version du certificat, 288855 (48,6%) l'ancienne. 57,7% des certificats papiers correspondaient à l'ancienne version.

Pour l'année 2019, 449 458 (75,1%) correspondaient à la nouvelle version du certificat, 149 249 (24,9%) l'ancienne. 30,8% des certificats papiers correspondaient à l'ancienne version.

## 2.4 Volets médicaux complémentaires (VMC)

Un volet médical complémentaire doit être rédigé en cas d'investigations médicales ou médico-légales *post-mortem*. Il est identique au volet médical initial et est complété sur la plateforme de certification électronique. Il se substitue au volet médical initial s'il apporte davantage d'informations concernant les causes du décès. Les VMC sont collectés à partir du 1<sup>er</sup> janvier 2018.

En 2018, le CépiDc a pu prendre en compte 1841 VMC dont la quasi-totalité (1 796) étaient issus de l'Institut médico-légal (IML) de Paris. En effet, cet IML a transmis l'ensemble de ces données issues de son système d'information directement au CépiDc. Les VMC de l'IML de Paris n'ont pas le même format que le volet médical initial pour la description des causes de décès car l'IML renseigne directement le code CIM de la cause initiale du décès telle que le médecin légiste la définit. Les autres 45 VMC pris en compte ont été collectés via l'application de certification électronique.

En 2019, le CépiDc a pris en compte 2 609 VMC dont 2466 issus de l'IML de Paris.

En conclusion, il existe une disparité géographique sur la collecte des VMC car leur complétion reste très rare dans les IML de province.

## 2.5 Type de certificat

En 2018, sur l'ensemble des volets médicaux reçus et conservés au CépiDc, 2 279 était des volets médicaux néonataux, dont un quart rédigé électroniquement. 363 de ces volets médicaux ont été utilisés pour des décès de plus de 28 jours. Inversement, 10 décès chez des bébés de moins de 28 jours ont été certifiés en utilisant un modèle général.

En 2019, 2 273 étaient des volets médicaux néonataux, dont 31% rédigés électroniquement. Et, 337 de ces volets médicaux ont été utilisés pour des décès de plus de 28 jours alors que seulement 5 décès chez des bébés de moins de 28 jours ont été certifiés sur un modèle général.

## 3 Codage

### 3.1 Contexte et stratégie

Pour produire la base sur les causes médicales des décès 2018 et 2019 et rattraper le retard de production accumulé, le CépiDc a mis à jour son processus de codage en intégrant un mode de codage impliquant des algorithmes de deep learning entraînés sur les millions de certificats de décès déjà codés. Ce processus a permis de produire une base de données provisoires en décembre 2022 avant l'intervention du codage manuel qui a eu lieu entre février et juin 2023.

Alors que deux modes de codage cohabitaient jusque-là (codage automatique sur batch avec le système expert Iris/Muse et codage manuel assisté par Iris/Muse pour les certificats non codés automatiquement par batch), un troisième mode de codage automatique (IA) a été utilisé. Ces algorithmes de deep learning peuvent intervenir aux deux étapes de codage (codage des expressions et identification de la cause initiale). Ils interviennent aussi pour cibler certains certificats envoyés en reprise manuelle. Ce paragraphe permet de décrire le processus de codage mis en place pour produire les données 2018 et 2019 en intégrant ce nouveau mode de codage. Il décrit également l'adaptation du processus de vérification usuellement réalisé dans une campagne de production traditionnelle.

#### 3.1.1 Modes de codage

Au total, pour les décès 2018 et 2019, le CépiDc a eu recours à trois modes de codage pour coder les causes multiples et déterminer la cause initiale :

- Le codage automatique à l'aide du système expert Iris/Muse (Batch) : 63% en 2018, 64% en 2019)
  - o Version du logiciel Iris : 5.5 en 2018 / 5.6 en 2019
  - o Version du moteur Muse utilisé dans Iris : 2.6 en 2018 / 2.7.1 en 2019
- Le codage manuel assisté par Iris/Muse (3%)
- Une prédiction par un algorithme de deep learning, possiblement combinée au système expert Iris/Muse: 34% en 2018, 33% en 2019. Ce mode de codage est détaillé dans le document *Codage des causes de décès de 2018 et 2019 en CIM10 - Approche combinant deep learning, système expert et codage manuel ciblé* », document de travail du CépiDc n°2, septembre 2023 ([disponible sur le site internet du CépiDc](#)).

Tableau 2. Répartition des modes de codage des données 2018 et 2019 (données finales en comparaison des données provisoires diffusées à l'hiver 2022-2023) hors volets médicaux non reçus

Année / Type de codage	Manuel	Codage prédit par algo IA	Codage automatique IRIS/MUSE	Total
2018 (Nbre)	18142	200217	376305	594664
2018 (%)	3%	34%	63%	100%
2018 provisoire - %	0%	37%	63%	100%
2019 (Nbre)	18805	196291	383611	598707
2019 (%)	3%	33%	64%	100%
2019 provisoire - %	0%	38%	62%	100%

### 3.1.2 Échantillons de codage manuel

Le codage manuel des certificats non codés par batch a été ciblé selon différentes catégories de certificats. Ceux-ci ont été rassemblés dans ce que l'on appelle par la suite des échantillons. Dans une campagne classique où tous les certificats non codés en batch étaient codés manuellement, les certificats étaient traités par les codeurs selon l'origine (lots papiers ou électroniques) et par ordre d'intégration en base de données (chronologique). C'est ce qu'on appelle le codage manuel du "tout venant". La transition vers un codage manuel par échantillon a pu avoir un impact sur la façon de coder, notamment quand un échantillon donné se concentre sur certaines catégories de causes de décès.

Pour la production 2018-2019, le codage manuel couvre :

- L'échantillon démographique permanent (EDP<sup>1</sup>) non codé en batch par Iris/Muse: 9892 certificats en 2018 ; 9705 en 2019
- Une partie des décès "sensibles". Il s'agit de types de décès pour lesquels le CépiDc doit garantir l'exactitude du codage pour des raisons de santé publique (VIH, morts maternelles, décès d'enfants). Le codage manuel concerne les certificats non codés par Iris/Muse : 3272

<sup>1</sup> Panel sociodémographique longitudinal de 4% de la population géré par l'insee : <https://www.insee.fr/fr/metadonnees/source/serie/s1166>

(année 2018) ; 3171 (année 2019). Les décès sensibles codés par batch par Iris/Muse sont quant à eux vérifiés par l'équipe de codage (cf processus de vérification 3.1.3). Certains de ces certificats ont été codés manuellement dans le cadre de l'EDP.

Tableau 3 Description et effectifs des décès identifiés comme sensibles en 2018 et 2019 hors volets médicaux non reçus

Catégorie de décès sensibles en 2018	Nbre de certificats
Moins de 28 jours (0-27 jours inclus)	1926
Au moins 28 jours – Moins de 15 ans	1975
Morts Maternelles	119
SIDA/VIH	644

Catégorie de décès sensibles en 2019	Nbre de certificats
Moins de 28 jours (0-27 jours inclus)	1941
Au moins 28 jours – Moins de 15 ans (dont morts violentes ou spécifiques* ou non codés automatiquement)	1918 (dont 1576)
Morts Maternelles	109
SIDA/VIH	643

Les décès sensibles correspondent en 2018 à toutes les mentions de SIDA/VIH sur le certificat, les morts maternelles ainsi que tous les décès de moins de 15 ans. Les décès sensibles correspondent en 2019 à toutes les mentions de SIDA/VIH sur le certificat, les morts maternelles, tous les décès de moins de 28 jours, les morts violentes de moins de 15 ans, les décès de moins de 15 ans avec mention d'un code spécifique (cf annexe 5.3), et tous les décès des moins de 15 ans non codés automatiquement par batch.

- Les échantillons issus des catégories les moins bien prédites par l'IA :

3116 (2018) ; 3202 (2019) en priorité 1 (échantillons de certificats constitués de façon à atteindre au moins 90% de cohérence de codage pour chaque catégorie de la short-list Eurostat)

Une partie des échantillons IA en priorité 2 - Les effectifs ciblés à ce titre sont 2646 (2018) et 2877 (2019) (choisis de façon à atteindre 92,5% de cohérence), et on en a effectivement repris manuellement 1717 pour 2018 et 2357 pour 2019, faute de moyens pour aller au-delà.

- Des lots de certificats ont été spécifiquement générés car la performance de la stratégie de codage menée jusqu'alors restait insuffisante pour certaines catégories de décès (sous-estimation) :
  - o en lien avec des homicides : dès lors qu'un modèle de deep learning remonte une cause de type homicide en cause initiale ou associée ou que les circonstances apparentes de décès mentionnent une atteinte à la vie d'autrui et que l'homicide



n'était pas retenu en cause initiale, 256 certificats codés manuellement en 2018 ; 244 en 2019

- en lien avec la pharmacodépendance et les tuberculoses : un modèle de deep learning remonte une cause de ce type en cause initiale et que la cause initiale retenue ne relève pas de la catégorie, soit pour la pharmacodépendance 46 en 2018 et 60 en 2019 et pour les tuberculoses 55 en 2018 et 89 en 2019

### 3.1.3 Processus de vérification

**Pour 2018 et 2019, les vérifications n'ont pas été réalisées de la même façon que lors d'une campagne traditionnelle de codage (comme en 2016, 2017 ou 2020).**

Traditionnellement, pour s'assurer de la qualité et de la cohérence du codage, des contrôles sur certaines catégories de certificats codés sont réalisés de façon exhaustive quel que soit le mode de codage. C'est le processus de vérification (cf document 'Statistiques sur les causes de décès de A à Z' [accessible sur le site internet du CépiDc.](#)). Avec l'introduction du troisième mode de codage par IA, le contexte de rattrapage contraignant, les moyens humains insuffisants et les délais de livraison des données, ce processus de vérification ne s'est pas déroulé comme traditionnellement. En outre, le codage manuel des échantillons issus des catégories les moins bien prédites de l'IA pouvaient contenir des certificats qui auraient été vérifiés dans une campagne traditionnelle.

Les paragraphes ci-dessous décrivent, dans la mesure du possible, les parties du processus qui ont été menées comme dans une campagne traditionnelle et celles qui ne l'ont pas été ou partiellement aux regards du changement de stratégie de codage manuel par échantillons ciblés.

#### 3.1.3.1 Cohérences

Une première série de vérifications consiste à vérifier les cohérences entre les codes de causes initiales et le sexe ou l'âge du défunt (cancer de la prostate chez une femme, pathologie hormonale à un âge trop ou pas assez avancé...). De plus, certains codes qui ne peuvent pas être utilisés en cause initiale de décès selon les recommandations doivent être modifiés. Ces vérifications ont toutes été réalisées. Cela a concerné 92 certificats en 2018, 131 en 2019.

#### 3.1.3.2 Décès sensibles à vérifier

L'OMS et les enjeux de santé publique mettent en évidence des décès sensibles pour lesquels il faut garantir un codage optimal selon les informations à disposition sur le certificat. Il s'agit de groupes de personnes ou pathologies spécifiques. Dans une campagne traditionnelle de codage, ces certificats étaient codés par le batch ou dans les lots "tout venant" par les codeurs. Puis, ils étaient tous vérifiés par des experts. Les certificats codés manuellement par un expert étaient considérés comme vérifiés au moment du codage. Dans la campagne de codage 2018-2019, une majorité de ces certificats étaient identifiés dans les décès sensibles tel que décrits en 3.1.2, dans l'EDP ou couverts par un échantillon de catégories les moins bien prédites de l'IA. Ils ont donc majoritairement été codés par batch ou par codage manuel assisté. Cependant, l'étape de vérification ultérieure par un expert, des certificats codés automatiquement ou manuellement par un non-expert n'a pas été menée comme dans une

campagne traditionnelle. On détaille ci-dessous les catégories pour lesquelles les vérifications ont été menées complètement à l'image d'une campagne traditionnelle de codage et celles pour lesquelles cela n'a pas été le cas.

Le tableau 4 présente les effectifs de décès inclus dans les décès sensibles tels que décrits en 3.1.2 et codés manuellement et le nombre d'entre eux dis « vérifiés » c'est-à-dire soit codé par un expert soit vérifié par un expert.

Tableau 4. Description des vérifications liées aux décès sensibles codés manuellement et nombre total de certificats associés. Parmi ces certificats, certains ont été vérifiés dans le cadre du codage manuel ciblé des « décès sensibles ».

Description	2018		2019	
	Total	Vérifiés	Total	Vérifiés
Décès d'enfants entre 0 et 27 jours (dont les enfants nés sans vie)	1926	658	1941	1941
Décès d'enfants entre 28 jours et 15 ans	1975	904		
Décès d'enfants entre 28 jours et 15 ans avec une mort violente ou pathologie spécifique (dont codes P)			1576	892
Certificats avec mention de SIDA/VIH*	679	639	672	643
Certificats de morts maternelles	119	119	109	109

\*Des certificats de SIDA/VIH ont été identifiés après le codage manuel ciblé décrit en 3.1.2. C'est le codage impliquant les algorithmes de deep-learning qui les a codés ainsi.

Lecture : parmi les 1 926 décès des moins de 28 jours en 2018, 658 ont été soit codés par un expert manuellement soit vérifiés par un expert. Les autres ont été codés manuellement par un nosologiste.

Pendant, le codage manuel ciblé dit « décès sensibles » ne couvrait pas l'ensemble des cas sur lesquels traditionnellement portaient une vérification. En effet, les vérifications d'une campagne traditionnelle couvraient aussi les intentions indéterminées, les mentions de suicide et d'homicide en causes associées (Tableau 5). Parmi eux, il se trouve que certains ont été codés manuellement dans le cadre de la campagne de codage manuel ciblé par l'équipe de codage. Par ailleurs, les certificats avec mention d'homicide en cause associée non retenue en cause initiale issus du codage impliquant les algorithmes de deep learning ont été codés manuellement (3.1.2). En revanche, cela n'a pas été le cas pour les mentions de suicide en cause associée.

Tableau 5. Description des catégories habituellement vérifiées et nombre total de certificats associés. Parmi ces certificats, certains ont été codés manuellement mais les vérifications n'ont pas été réalisées.

Description	2018		2019	
	Total	Codage Manuel	Total	Codage Manuel
Les certificats ayant une cause initiale de la catégorie intention indéterminée (commençant par Y1 ou Y2 ou Y3[0-4])*	1195	246	1313	305
Les certificats avec une mention de suicide mais non retenu en cause initiale*	666	24	698	25

Les certificats avec une mention d’homicide mais non retenu en cause initiale*	375	263	336	249
--	-----	-----	-----	-----

\*Les certificats codés manuellement l’ont été dans le cadre de la reprise des catégories les moins bien prédites de l’IA ou de l’EDP (3.1.2).

Lecture : parmi les 666 certificats avec une mention de suicide mais non retenue en cause initiale, 24 ont été codés manuellement en 2018 dans le cadre des différents échantillons décrit en 3.1.2. Les autres certificats auraient dû être vérifiés pour s’assurer de la qualité du codage et comprendre pourquoi le suicide n’est pas retenu (le cas échéant, corriger le codage).

### 3.1.3.3 Difficultés de codage connues liées au logiciel Iris

Le logiciel (système expert IRIS/MUSE) qui permet le codage par batch, le codage manuel semi-assisté et une partie du codage réalisé avec des algorithmes de deep learning a des lacunes identifiées. Il était donc habituel, lors des campagnes précédentes, de vérifier certains certificats parmi la liste ci-dessous.

**Pour 2018 et 2019, ces vérifications n’ont pas été faites.** Elles auraient concerné 2616 certificats en 2018 et 2187 en 2019 au total.

Tableau 6. Description des vérifications liées au système expert et nombre total de certificats associés. Parmi ces certificats, certains ont été codés manuellement mais les vérifications n’ont pas été réalisées.

Description	2018		2019	
	Total	Codage Manuel	Total	Codage Manuel
Cause initiale en Y avec un code T sur le certificat : dans ces cas-là, le plus souvent une autre pathologie doit être en CI	1012	265	923	219
Vérification des certificats avec un code D00-D48 en CI alors qu’il y a mention de cancer	435	10	443	14
Traitement du "ou" si mention de mort violente à côté du "ou"	388	98	370	102
Code d'affection hématologique (D50-D89) en cause initiale alors qu’il existe un code C sur le certificat	372	84	179	29
Notion de surdosage AVK en partie 1	251	12	249	20
D611 en CI et absence de cancer dans le certificat	143	52	0	0
Cause initiale à coder en C80.9 si C34.9 ou un C41.1 en cause associée	15	0	23	4
<i>Mention de "sur" suivi de certains codes laissant à penser que le « sur » ne doit pas être considéré comme un lien de causalité, avec impact potentiel sur la CI</i>	<i>cette vérification n’a pas été générée car les prédictions impliquant des algorithmes de deep-learning ne permettent pas de les identifier.</i>			

Lecture : 435 certificats avec un code D00-D48 en CI alors qu’il y a mention de cancer sont habituellement vérifiés car le système expert contient une anomalie qui parfois n’identifie pas la bonne cause initiale (ici le cancer est le plus probable). En 2018, 10 de ces certificats se trouvaient dispersés dans les échantillons de codage manuel (assisté par le système expert). Les codeurs n’ont pas réalisé

la vérification spécifique de cette problématique mais le codage manuel a pu prendre en compte cette vigilance.

#### 3.1.3.4 Vérification de la bonne application des nouvelles règles de 2016

L'OMS a introduit des changements dans les règles de sélection de la cause initiale depuis 2016. Selon la complexité de ces changements, des vérifications sont encore réalisées pour s'assurer de la bonne compréhension et application de ces règles. En 2018 et 2019, cela ne concernait plus que les hépatites virales et le diabète.

**Ces vérifications n'ont pas été faites réalisées.** Elles concernaient 449 certificats en 2018 et 471 certificats en 2019.

Tableau 7. Description des vérifications liées aux nouvelles règles et nombre total de certificats associés. Parmi ces certificats, certains ont été codés manuellement mais n'ont pas été vérifiés.

Description	2018		2019	
	Total	Codage Manuel	Total	Codage Manuel
Les hépatites virales ne peuvent plus être dues à d'autres pathologies : les certificats avec mention d'hépatites virales non retenues en cause initiale	326	18	343	37
Le diabète peut désormais être dû à une liste de pathologies restreintes	123	3	128	4

#### 3.1.3.5 Vérifications liées aux évolutions du dictionnaire (démarche des choix de code)

A partir de 2018, le dictionnaire a été mis à jour, notamment les libellés faisant l'objet de "choix de code". On définit un "choix de code" comme étant un libellé qui peut être codé de deux façons différentes selon le contexte du certificat. Cela nécessitait systématiquement un regard humain alors que dans certains cas, le choix était toujours le même ou les autres n'étaient pas pertinents. Des suppressions de choix de codes au profit d'un code "préféré" ont été réalisées et ont permis de diminuer les certificats inutilement codés manuellement. Cependant, certaines situations peuvent laisser le doute et des vérifications ont donc été définies.

Par exemple, le libellé "épuisement" pouvait être codé soit en R53 (malaise, fatigue) soit en T73.3 (surmenage lié à effort intensif) soit en X50 (surmenage lié à effort répété). Il a été décidé de supprimer le choix de code pour ce libellé et de permettre le codage automatique en R53. Cependant, il a été décidé de vérifier les certificats qui mentionnent un épuisement chez les moins de 75 ans car c'est atypique et pourrait correspondre à un effort intensif ou sportif. Le contexte du certificat permet au codeur, lors de sa vérification de statuer et d'éventuellement recoder l'épuisement correctement.

Le tableau ci-dessous décrit les vérifications définies suite à cette démarche de suppression de choix de code dans le dictionnaire. Celles-ci n'ont pas été réalisées sur 2018 et 2019.

Tableau 8. Description des vérifications liées à la démarche des choix de code et nombre total de certificats associés. Parmi ces certificats, certains ont été codés manuellement mais n'ont pas été vérifiés.

Description	2018		2019	
	Total	Codage Manuel	Total	Codage Manuel
S'il y a une notion de causes externes dans un certificat avec mention d'hémorragie intracérébrale, le codage sera S068 et non I6199	855	44	777	47
On vérifie que le code J96.1 est approprié pour les certificats avec un libellé « IRC » et une notion de diabète ou certaines pathologies rénales/urinaires.	556	14	621	23
S'il y a une notion de traumatisme dans un certificat avec mention d'hémorragie interne, le code R5809 ne sera pas forcément approprié	482	41	441	29
S'il y a une notion de cause externe dans un certificat avec mention d'alcool, on vérifie que le code F10.2 est approprié	368	26	356	28
On vérifie que le code I50.1 est approprié pour le libellé "OAP" dans les certificats avec mention d'"IR" et "IC" (potentiellement à reprendre en J81)	254	9	259	28
S'il y a une notion d'acte diagnostique et thérapeutique invasif dans un certificat avec mention de choc hémorragique ou hypovolémique, on vérifie que le code R57.1 est approprié (potentiellement à reprendre en T179)	233	34	234	33
On vérifie que le code J96.0 est approprié pour les certificats avec un libellé « IRA » et une notion de diabète ou certaines pathologies rénales/urinaires.	133	7	143	9
Libellé « épuisement » chez une personne de moins de 75 ans est souvent dû à un effort sportif. On vérifie que le code R53 est approprié chez cette catégorie	129	5	125	7
S'il y a une notion de traumatisme ou cause externe dans un certificat avec mention de rhabdomyolyse, on vérifie que le code M62.8 est approprié.	70	5	82	6

## 3.2 Evolutions de codage

### 3.2.1 Nouveautés relatives aux recommandations OMS

La nouvelle version du certificat introduite en 2018 met en évidence un processus morbide plus précis. De plus, les « circonstances apparentes de décès » collectées sous forme de cases à cocher facilitent le codage surtout lorsqu'il y a présence de causes externes.

Malgré les recommandations de l’OMS, la France a fait le choix de conserver deux certificats bien distincts en fonction de l’âge du défunt. Il y a donc eu l’apparition d’un nouveau certificat général (cf 5.1) et d’un nouveau certificat néonatal (cf 5.2). Ce certificat néonatal contient les causes obstétricales ou maternelles qui permettent un codage plus précis mais qui sont collectées dans un format différent que celui recommandé par l’OMS. Cela rend compliquée l’application des règles de codage puisqu’il faut adapter ces informations au format « général ».

### 3.2.2 Précisions dans l’application des règles de l’OMS

Les règles de codage décrites dans le Volume 2 de la CIM-10 depuis 2016, sont appliquées à partir de l’année 2017 et continuent sur 2018. Comme sur 2017, le code D68.3 continue à remplacer le code Y44.2 en cause initiale. En effet, pendant longtemps, l’anticoagulation en cause initiale était codée en Y44.2 mais un code plus précis existe dans le Chapitre III Maladies du sang et des organes hématopoïétiques et certains troubles impliquant le mécanisme immunitaire (D50-D89).

L’OMS recommande depuis 2016, de ne pas retenir les codes peu informatifs I46.0 (arrêt cardiaque réanimé) et I46.9 (arrêt cardiaque sans précision) en causes initiales. Dans les données du CépiDc, ces dernières deviennent systématiquement des R99 à partir de 2019. En 2018, 4 711 certificats ont une CI en I469 ou I460 qui aurait dû être R99.

### 3.2.3 Mises à jour du dictionnaire des expressions nosologiques utilisé par Iris Muse

Dans le cadre du codage, il arrive de mettre en évidence des codes imprécis, faux ou obsolètes pour certaines pathologies dans le dictionnaire des expressions nosologiques utilisé par Iris.

Ci-dessous la liste des libellés pour lesquels les codes CIM10 ont été mis à jour dans le dictionnaire. Jusque-là, plusieurs codes étaient possibles mais un seul était correct. Pour ces libellés, les codes supprimés ne seront plus rencontrés à partir de 2018. La qualité du codage a donc été améliorée, mais cela peut induire des ruptures de séries avec les années précédentes.

Libellé	Code correct
Inhalation	T179
Pacemaker	I499
Décompensation respiratoire	J969
Cirrhose hépatique	K746
Asphyxie	R090
Encéphalopathie	G934
Hématurie	N029
Coma hyperosmolaire	E140
GIR 1	R263
Syndrome Korsakoff	F04

## 4 Evolution de la base de données

### 4.1 Nouvelle version de certificats de décès

Avec la nouvelle version du certificat de décès, des nouvelles informations sont collectées et des nouvelles variables ont été ajoutées à la base de données :

- L'information sur la réalité de la date de décès (date de la mort VS date du constat du corps décédé) dans le certificat général
- Les circonstances apparentes de décès,
- Le lieu de survenue en cas de mort violente pour le certificat général
- L'information d'une mort subite ou de mort inattendue du nourrisson
- La contribution de la grossesse au décès pour le certificat général
- Le lieu du décès a été ajouté au certificat néonatal

D'autres informations ont vu leur libellé et/ou leur modalités de réponse évoluer :

- La grossesse au moment du décès ou dans l'année qui a précédée est davantage détaillée dans le certificat général et permet de supprimer une information mal renseignée en texte libre sur l'intervalle entre la fin de la grossesse et le décès.
- L'accident de travail est reformulé en "le décès a-t-il eu lieu pendant une activité professionnelle ?" dans le certificat général
- La recherche de la cause de décès vient remplacer la notion d'autopsie et il est possible de préciser le type de recherche (médicale ou médico-légale).
- L'établissement pénitentiaire est une modalité de plus pour le lieu du décès dans le certificat général

Les détails sur les variables et leurs modalités de réponse sont dans l'annexe 5.4.

### 4.2 Nouveau mode de codage

Dans le cadre des évolutions du processus de codage impliquant de nouveaux outils, le CépiDc a mis en place des informations complémentaires disponibles en base de données :

- Le type de codage qui correspond à la méthode utilisée : codage manuel, automatique par système expert Iris/Muse ou autre méthode de codage (notamment impliquant des algorithmes de deep-learning, lesquels sont détaillés).
- L'indicateur de confiance associé au codage automatique impliquant des algorithmes de deep-learning. Cet indicateur est celui utilisé pour cibler les certificats à coder manuellement (cf documentation sur le codage 2018-2019 impliquant des algorithmes de deep-learning).

Le détail sur ces deux nouvelles variables et leurs modalités de réponse sont dans l'annexe 5.4





## 5.2 Version 2017 du volet médical du certificat de décès néonatal (moins de 28 jours)

VOLET MÉDICAL. À remplir et à clore par le médecin ayant constaté le décès – Renseignements confidentiels et anonymes (* Instructions en annexe)			
<b>INFORMATIONS RELATIVES À L'ENFANT</b>			
Commune de décès :		Code postal :	
Commune de domicile :		Code postal :	
Date et heure de décès :		Date et heure de naissance* :	
à h		à h	
Sexe :			
<input type="checkbox"/> masculin			
<input type="checkbox"/> féminin			
<input type="checkbox"/> indéterminé			
Appar à 1 minute :		Âge gestationnel en semaines révolues d'aménorrhée :	
		Poids de naissance en grammes :	
<b>INFORMATIONS RELATIVES À L'ACCOUCHEMENT</b>		<b>INFORMATIONS RELATIVES AUX PARENTS</b> (inscrire le code approprié)	
Naissance : 1. unique 2. gémellaire 3. triple 4. quadruple 5. quintuple <input type="checkbox"/> Numéro d'ordre de l'enfant si grossesse multiple : <input type="checkbox"/> Lieu d'accouchement : 1. établissement de santé 2. domicile 3. autre <input type="checkbox"/> Présentation : 1. sommet 2. autre céphalique 3. siège 4. autre <input type="checkbox"/> Début du travail : 1. spontané 2. déclenché 3. césarienne avant travail <input type="checkbox"/> Mode d'accouchement* : 1. voie basse non instrumentale <input type="checkbox"/> 2. extraction instrumentale par voie basse 3. césarienne <input type="checkbox"/> Transfert ou hospitalisation particulière *de l'enfant : 1. oui 2. non <input type="checkbox"/>		MÈRE <span style="float: right;">Année de naissance : <input type="text"/></span> Nationalité (en clair) : <input type="text"/> Profession* (en clair) : <input type="text"/> exercée pendant la grossesse : 1. oui 2. non 3. chômage 4. autre situation <input type="checkbox"/> État matrimonial : 1. célibataire 2. mariée 3. veuve 4. divorcée <input type="checkbox"/> La mère vit-elle en couple ? 1. oui 2. non <input type="checkbox"/> Nombre total de grossesses, y compris grossesse pour cet enfant : <input type="text"/> Nombre total d'accouchements, y compris accouchement pour cet enfant* : <input type="text"/>	
		PÈRE <span style="float: right;">Année de naissance : <input type="text"/></span> Profession* (en clair) : <input type="text"/> exercée pendant la grossesse : 1. oui 2. non 3. chômage 4. autre situation <input type="checkbox"/>	
<b>CAUSES DU DÉCÈS</b> (*Lire les instructions de remplissage en annexe)			
<b>CAUSE FŒTALE OU NÉONATALE*</b> déterminante de la mort – Affection(s) morbide(s) ayant directement provoqué le décès. Il s'agit de la maladie, du traumatisme, de l'intoxication, de la complication ayant entraîné la mort (et non du mécanisme de décès comme une syncope, un arrêt cardiaque...). a) <input type="text"/> due à ou consécutive à : b) <input type="text"/> due à ou consécutive à : c) <input type="text"/> Autre(s) cause(s) fœtale(s) ou néonatale(s) associée(s) : <input type="text"/>			
<b>CAUSE OBSTÉTRICALE OU MATERNELLE*</b> déterminante de la mort : <input type="text"/> Autre(s) cause(s) obstétricale(s) ou maternelle(s) associée(s)* : <input type="text"/>			
<b>INFORMATIONS COMPLÉMENTAIRES</b> (cocher la case appropriée pour chaque point – *Lire les instructions de remplissage en annexe)			
<b>LIEU DU DÉCÈS</b> <input type="checkbox"/> Domicile (du défunt ou autre) <input type="checkbox"/> Établissement de santé public <input type="checkbox"/> Voie publique <input type="checkbox"/> Établissement de santé privé <input type="checkbox"/> Autre lieu ou indéterminé		<b>RECHERCHE DE LA CAUSE DU DÉCÈS*</b> Une recherche de la cause du décès a-t-elle été demandée ? <input type="checkbox"/> oui, recherche médicale <input type="checkbox"/> oui, recherche médico-légale <input type="checkbox"/> non Si oui, un volet médical complémentaire sera établi séparément par le médecin ayant réalisé le diagnostic des causes de décès.	
<b>MORT INATTENDUE DU NOURRISSON</b> S'agit-il d'un décès brutal et inattendu* ? <input type="checkbox"/> oui <input type="checkbox"/> non <input type="checkbox"/> ne sait pas * décès non traumatique du nourrisson avec mode de survenue brutal (en moins d'une heure ou probablement) et inattendu.		<b>SIGNATURE</b> Nom lisible et cachet obligatoire du médecin <input type="text"/>	
<b>CIRCONSTANCES APPARENTES DU DÉCÈS</b> <input type="checkbox"/> Mort naturelle <input type="checkbox"/> Faits de guerre <input type="checkbox"/> Accident <input type="checkbox"/> Complications de soins médicaux, chirurgicaux <input type="checkbox"/> Atteinte à la vie de l'enfant <input type="checkbox"/> Investigations en cours <input type="checkbox"/> Indétectées			
Ce volet n'est destiné qu'aux personnes autorisées pour des motifs de santé publique (cf article L. 2223-4) du Code général des collectivités territoriales).			
Le certificat peut être saisi électroniquement à l'adresse suivante <a href="http://mc.cerdds.inserm.fr">http://mc.cerdds.inserm.fr</a>			

### 5.3 Certificats des décès entre 28 jours et 15 ans identifiés dans les décès sensibles hors morts violentes

Descriptif	Code en CI ou en cause associée (Chapitre ou Intervalle de codes)	Exceptions (non considéré comme sensible à vérifier)
Maladie du système nerveux à l'exception de ...	G%	G12%, G40%, G41%, G70%, G71%, G72%, G80%, G93%
Troubles mentaux et du comportement	F%	
Maladie de l'appareil respiratoire à l'exception de ...	J%	J09%, J10%, J11%, J12%, J21%, J35%, J45%, J46%, J840
Maladie de l'appareil digestif à l'exception de ...	K%	K35%, K65%
Maladie de la peau et du tissu cellulaire sous-cutané	L%	
Tumeur à l'exception de ...	C01-C98	C222, C40%, C41%, C49%, C62%, C64%, C71%, C72%, C91%, C92%, C93%, C94%, C95%

## 5.4 Dictionnaire de variables dans le SNDS

Table Ident

Colonne	Nom de la variable	Formats	Modalités des valeurs SNDS	Commentaires
1	Identifiant IdDécès	varchar (64)		
2	Version du certificat	Num	1 = version de 1997 2 = version de 2017	La version 2017 apparaît à partir de 2017 (expérimentation)
3	Statut de traitement	Num	1 = Codé 2 = En cours	Les données finalisées envoyées dans le SNDS sont toutes à 1
4	Type de certificat	Num	1 = adulte 2 = néonatal	Variable de mauvaise qualité ; pour identifier le type de certificat, regarder l'âge du défunt et la présence ou non de valeurs dans les variables spécifiques du certificat néonatal
5	Type de support	Num	1 = électronique 2 = papier	
6	Type de volet	Num	1 = initial 2 = Complémentaire	Les volets médicaux complémentaires apparaissent à partir de 2018
7	Département de décès	Varchar (3)		2 caractères pour métropole, 3 pour DOM TOM
8	Commune de décès	Varchar (3)		3 caractères pour métropole, 2 pour DOM TOM
9	Département de domicile	Varchar (3)		2 caractères pour métropole, 3 pour DOM TOM
10	Commune de domicile	Varchar (3)		3 caractères pour métropole, 2 pour DOM TOM
11	Date du décès	Date		

12	Lieu du décès	num	0= Non renseigné 1 = Domicile 2 = Etablissement public de santé 3 = établissement privé de santé 4 = EHPAD, maison de retraite, 5 = voie publique 6 = autre lieu ou indéterminé 7=établissement pénitentiaire (certificat de 2017 uniquement)	
13	Année de naissance	Car (4)		
14	Mois de naissance	Car (2)		
15	Sexe	num	1 = masculin 2 = féminin 9 = indéterminé	
16	Activité professionnelle	num	1 = Retraité 2 = Inactif autre que retraité 3 = actif	Variable de l'INSEE
17	Profession et catégorie socio-professionnelle	varchar (2)	Nomenclature PCS INSEE	Variable de l'INSEE
18	Etat matrimonial	num	1 = célibataire 2 = marié 3 = veuf 4 = divorcé	Variable de l'INSEE
19	Cause initiale du décès	varchar (4)	CIM 10	

20	Recherche de la cause de décès	num	<p>1 = non</p> <p>2 =oui, résultats disponibles</p> <p>3 =oui, résultats non disponibles</p> <p>4 =oui recherche médicale (certificat de 2017 uniquement),</p> <p>5 = oui recherche médico-légale (certificat de 2017 uniquement)</p>	<p>Valeur vide de 2006 à 2009, question n'était pas posée avant 2010.</p> <p>Valeur 2 et 3 uniquement pour les versions 1997</p> <p>Valeur 4 et 5 uniquement pour les nouveaux certificats à partir de 2017</p>
21	<p>Grossesse</p> <p>"Le décès est-il survenu pendant une grossesse (ou moins d'un an après) " (ancienne version) ?</p> <p>La femme décédée était-elle enceinte ?" (nouvelle version)</p>	num	<p>1 = non, pas au cours de l'année précédant le décès</p> <p>2 = Pas au moment du décès mais grossesse terminée depuis 42 jours au moins</p> <p>3 = Pas au moment du décès mais grossesse terminée depuis plus de 42 jours et moins de 1 an</p> <p>4 = Oui, au moment du décès</p> <p>5 = Ne sait pas</p>	<p>Cette variable est vide avant 2015</p> <p>Libellé de la question et modalités de réponses ont changé entre la version 1997 et 2017 mais la variable reste la même</p>
22	Délai entre fin de grossesse et décès	Varchar(4)	<p>mois + jour, ex : 0502</p> <p>= 5 mois + 2 jours après la date de fin de grossesse</p>	<p>Variable mal renseignée, de mauvaise qualité et uniquement sur la version 1997.</p>
23	<p>Ancien certif : Est-ce un Accident de travail ?</p> <p>Nouveau certif : Est-ce pendant une activité professionnelle ?</p>	num	<p>1 = oui</p> <p>2 = Non</p> <p>3 = sans précision (ancien) ou Ne sait pas (nouveau)</p>	<p>Libellé de la question et modalités de réponses ont changé mais la variable reste la même</p>
24	Apgar à une minute	num	<p>0 à 10 (score d'Apgar, 0 = mort apparente, 10 = état optimal)</p>	<p>Variable spécifique du certificat néonatal</p>
25	Âge gestationnel en semaines révolues d'aménorrhée	num	en semaines	variable spécifique du certificat néonatal
26	Poids de naissance en grammes	num	en g	variable spécifique du certificat néonatal

27	Type de naissance	num	1 = Unique 2 = Gémellaire 3 = Triple 4 = Quadruple 5 = Quintuple	variable spécifique du certificat néonatal
28	N° d'ordre de l'enfant si grossesse multiple	num	1 à 5	variable spécifique du certificat néonatal
29	Lieu d'accouchement	num	1 = Etablissement de sante 2 = Domicile 3 = Autres	variable spécifique du certificat néonatal
30	Présentation de l'enfant	num	1 = Sommet 2 = Autres céphaliques 3 = Siège 4 = Autres	variable spécifique du certificat néonatal
31	Début du travail	num	1 = Spontané 2 = Déclenché 3 = Césarienne	variable spécifique du certificat néonatal
32	Mode d'accouchement	num	1 = Voie basse 2 = Extraction 3 = Césarienne	variable spécifique du certificat néonatal
33	Transfert ou hospitalisation particulière de l'enfant	num	1 = Oui 2 = Non	variable spécifique du certificat néonatal
34	Année de naissance de la mère	Varchar(4)		variable spécifique du certificat néonatal
35	Activité professionnelle de la mère	num	1 = En activité 2 = Non au chômage 3 = Autres 4 = Au chômage 5 = Non	variable spécifique du certificat néonatal 2006 à 2015 uniquement les 3 1ere modalités, les autres modalités arrivent avec le nouveau certificat néonatal en 2017
36	Profession de la mère exercée pendant la grossesse	Varchar (50)	libellé de la profession	variable spécifique du certificat néonatal

37	Etat matrimonial de la mère	num	1 = célibataire 2 = mariée 3 = veuve 4 = divorcée	variable spécifique du certificat néonatal
38	La mère vit elle en couple	num	1 = Oui 2 = Non	variable spécifique du certificat néonatal
39	Nombre total de grossesses	num		variable spécifique du certificat néonatal
40	Nombre total d'accouchements	num		variable spécifique du certificat néonatal
41	Activité professionnelle du père		1 = En activité 2 = Non au chômage 3 = Autres 4 = Au chômage 5 = Non	variable spécifique du certificat néonatal 2006 à 2015 uniquement les 3 1ere modalités, les autres modalités arrivent avec le nouveau certificat
42	Profession du père exercée pendant la grossesse		libellé de la profession	
43	Mort subite	num	1 = Ne sait pas 2 = Oui 3 = Non	Uniquement dans la version 2017
44	Circonstance apparente du décès	num	1 = Indéterminé 2 = Mort naturelle 3 = Accident 4 = Suicide 5 = Atteinte à la vie du défunt 6 = Fait de guerre 7 = Investigations en cours 8 = Complications de soins 9 = Atteinte à la vie de l'enfant (certificat néonatal uniquement) ;	Uniquement dans la version 2017

45	Codification du lieu si mort violente	num	1 = Autre lieu ou indéterminé 2 = Domicile 3 = Etablissement accueillant du public 4 = Exploitation agricole 5 = Lieu de sport 6 = Commerce 7 = Voie publique 8 = Local industriel - chantier	Uniquement dans la version 2017
46	Fiabilité de la date de décès	num	1 = date réelle 2 = Hypothèse	Uniquement dans la version 2017
47	Mort inattendue du nourrisson	num	1 = Ne sait pas 2 = Oui 3 = Non	variable spécifique du certificat néonatal Uniquement dans la version 2017
48	Groupe d'âge mortalité néonatale		1 = < 7 jours 2 = 7 jours <= âge < 28 jours 3 = 28 jours <= âge <= 365 jours	
49	La grossesse a-t-elle contribué au décès ?	num	1 = Oui 2 = Non 3 = Ne sait pas	Uniquement dans la version 2017
50	CauseInitialeTypeCode	num	1 = CIM 9 2 = CIM 10 3 = CIM 11	
51	TypeCodage	num	1=codage Iris Automatique 2= manuel 3= K5Iris 4=K4Iris 5=K5 6=K4	Uniquement à partir de 2018 Décrit le mode de codage utilisé. Les types de codage 3 à 6 impliquent des algorithmes de deep-learning
52	Score de confiance	num	Entre 0 et 1 avec 3 décimales	Uniquement à partir de 2018 Indicateur de confiance issu de la prédiction impliquant des algorithmes de deep learning permettant de cibler le codage manuel



Table des causes

Colonne	Nom de la variable	Formats	Modalités des valeurs SNDS	Commentaires
1	identifiant IdDécès	Varchar (64)		
2	N° de ligne de la cause sur le certificat	numérique	1 à 6	Numéro de la ligne de cause
3	rang de la cause	numérique	1 à 40	rang de la cause sur la ligne si plusieurs causes par ligne
4	libellé de la cause	Varchar (400)	texte de la cause	Pour les certificats codés avec algorithme de deep-learning (typecodage >=3) le texte de la ligne correspondra au texte brut présent sur le certificat sans découpage par rang
5	code CIM de la cause	Varchar (4)	Tables ir_cci_v* et ir_cim_v*	Ce champ va évoluer avec la CIM 11 et passera en varchar 6
6	TypeCodage			Variables non informatives à l'échelle de la ligne de cause, contenu vide.
7	Score de confiance			Variables non informatives à l'échelle de la ligne de cause, contenu vide.

## 5.5 Liste des tableaux

<i>Tableau 1. Exhaustivité de la collecte des volets médicaux (VM) pour les décès 2018 et 2019</i>	4
<i>Tableau 2. Répartition des modes de codage des données 2018 et 2019 (données finales en comparaison des données provisoires diffusées à l'hiver 2022-2023) hors volets médicaux non reçus</i>	7
<i>Tableau 3 Description et effectifs des décès identifiés comme sensibles en 2018 et 2019 hors volets médicaux non reçus</i>	8
<i>Tableau 4. Description des vérifications liées aux décès sensibles codés manuellement et nombre total de certificats associés. Parmi ces certificats, certains ont été vérifiés dans le cadre du codage manuel ciblé des « décès sensibles ».</i>	10
<i>Tableau 5. Description des catégories habituellement vérifiées et nombre total de certificats associés. Parmi ces certificats, certains ont été codés manuellement mais les vérifications n'ont pas été réalisées.</i>	10
<i>Tableau 6. Description des vérifications liées au système expert et nombre total de certificats associés. Parmi ces certificats, certains ont été codés manuellement mais les vérifications n'ont pas été réalisées.</i>	11
<i>Tableau 7. Description des vérifications liées aux nouvelles règles et nombre total de certificats associés. Parmi ces certificats, certains ont été codés manuellement mais n'ont pas été vérifiés.</i>	12
<i>Tableau 8. Description des vérifications liées à la démarche des choix de code et nombre total de certificats associés. Parmi ces certificats, certains ont été codés manuellement mais n'ont pas été vérifiés.</i>	13