



# OSIRIS RWD

## Guide d'implémentation

06/05/2024

# PRÉAMBULE

OSIRIS est un projet qui a pour but de favoriser le partage des données et créer les conditions de l'interopérabilité en oncologie. OSIRIS RWD (pour Real World Data) est une adaptation aux données « en vie réelle » du modèle OSIRIS publié en 2021 <sup>(1)</sup>.

Le modèle OSIRIS RWD inclut un Minimum Data Set (MDS) de données cliniques, complété par un module génomique et un MDS de données concernant l'imagerie et la radiothérapie dans le cadre du programme MED-OSIRIS (Maintenance, Extension et Déploiement du modèle OSIRIS).

Le MDS clinique a été défini sur la base d'un consensus d'experts. Il est complété d'un guide d'implémentation qui a été élaboré en parallèle à la liste des items.

Un modèle de données est en cours de définition (modèle conceptuel, modèle logique et modèle physique) et l'objectif est une mise en œuvre opérationnelle dans une base de données « open source ». Le choix s'est porté sur PostgreSQL, système de gestion de base de données relationnelle (SGBDR).

En aval de la base de données OSIRIS RWD, il est prévu d'extraire et transformer ces données vers d'autres formats : OMOP (pour le partage) et FHIR (pour la communication et le retour au dossier médical).

A terme, un des objectifs du projet OSIRIS RWD est de permettre l'interopérabilité entre les systèmes d'information hospitaliers (SIH) et les entrepôts de données de santé (EDS) hospitaliers. L'interopérabilité s'entend sous tous ses aspects : technique, syntaxique, sémantique et juridique.

- Technique : les données sont indépendantes de l'infrastructure des systèmes et des protocoles de communication
- Syntaxique : le modèle de données, le format des tables et le nom de la variable sont documentés
- Sémantique : les données sont codées de manière standardisée et normalisée avec un même sens donné aux codes et à la description de la variable
- Juridique avec un même respect des règles et des politiques de confidentialité pour être conforme au RGPD.

Dans ce document sont définies les règles d'alimentation de la base de données qui constituent le guide d'implémentation. L'alimentation de la base de données et le contrôle de la qualité des données sont sous la responsabilité des établissements de santé et des gestionnaires des EDS hospitaliers.

---

<sup>1</sup> OSIRIS: A Minimum Data Set for Data Sharing and Interoperability in Oncology  
Julien Guerin, Yec'han Laizet, Vincent Le Texier and all. JCO Clin Cancer Inform 2021, 5:256-265.

Ce document a pour objet de détailler les items du MDS et de proposer les modalités d'alimentation qui vont de la source (le SIH) vers la base de données pour une alimentation automatique. Différents points sont abordés en fonction de la complexité des items :

- Définition de la variable : à quelle entité elle appartient (patient, facteur de risque lié au cancer), caractère unique ou multi-valuée.
- Choix des référentiels (Ex : CIM10, LOINC, SNOMED CT).
- Définir la source : dossier médical textuel ou donnée structurée, issu de quel type de compte-rendu du dossier médical en tenant compte des possibilités de structuration par des techniques de traitement automatique du langage.

## Descriptif du document

Le guide d'implémentation regroupe les recommandations qui ont émanées du groupe d'experts. Les recommandations ont été validées par consensus.

Chaque item sera décrit comme suit : nom de l'item, définition, nomenclature utilisée, valeurs attendues, règles, options et recommandations (ROR) pour l'alimentation automatique.

### Les principes « FAIR » doivent s'appliquer

Le modèle de données OSIRIS RWD répond aux principes dits FAIR (Findable, Accessible, Interoperable, Reusable) ; ces principes décrivent comment les données sont organisées pour être accessibles, comprises, échangeables et réutilisables.

- Le principe « Findable » a pour objectif de permettre le traitement de la donnée qui doit être comprise autant par des humains que par des systèmes informatiques. Ceci requiert une description des données et de tout ce qui sert à caractériser une donnée, comme sa définition et son format.
- Le principe « Accessible » nécessite de stocker durablement les données et les métadonnées et faciliter leur accès en tenant compte des conditions d'accès (ouvert ou restreint) et d'utilisation (licence).
- Le principe « Interoperable » signifie que la donnée est standardisée sur le plan national voire international.
- Le principe « Reusable » met en avant les caractéristiques qui rendent les données réutilisables à d'autres finalités que les soins c'est-à-dire pour la recherche, l'enseignement, l'innovation ou l'évaluation.

### Les 10 principes spécifiques au modèle OSIRIS RWD

Dans le choix des items et des règles de structuration, plusieurs principes ont été adoptés que l'on peut résumer en dix points :

1. **Utile**, l'information collectée doit répondre à des objectifs de soins, de recherche, d'évaluation de pratique. Le MDS exclut des données pour effectuer des études de marché, comparer des hôpitaux entre eux dans le cadre d'une éventuelle concurrence.
2. La donnée doit pouvoir être collectée en **automatique** à partir des dossiers médicaux et des multiples sources des systèmes d'information hospitalier.
3. La mesure de **qualité** des données d'une base alimentée à partir de données transformées est indispensable. Il est proposé la création de bases de données de référence avec une saisie manuelle (« gold standard ») pour mesurer la qualité des données extraites automatiquement ou structurées à partir de textes via des règles ou des systèmes de traitement automatique du langage.
4. **Implémentable** ce qui veut dire que l'item doit être bien défini. Toutes les données sont décrites afin que l'alimentation automatique se fasse sans ambiguïté, ce qui nécessite une description précise de chaque variable.
5. Dans le respect de la **réglementation** en vigueur (RGPD/CNIL). Le respect de la réglementation s'impose et doit permettre de gérer de façon identique les données identifiantes et non identifiantes, à charge pour l'établissement de gérer une table de correspondance.
6. Les recommandations doivent être **reproductibles** dans le temps (valable pour des dossiers médicaux récents et anciens) et dans l'espace (pour tout hôpital quel que soit son système d'information). Les établissements hospitaliers sont susceptibles de conserver des données pendant un temps long (plusieurs années). Malgré les modifications de prises en charge dues aux progrès dans les méthodes diagnostiques ou thérapeutiques, les recommandations du guide doivent s'appliquer pour fournir des résultats cohérents sur une longue période même si le SIH évolue dans le temps.
7. Privilégier les **données structurées** à la source quand sont disponibles dans les SIH, par rapport à une source de données non structurée.
8. Lorsqu'on utilise un référentiel, on l'utilise avec les propres **règles du référentiel**. Par exemple, la CIM-O-3, que nous utilisons pour gérer les codes morphologiques de la tumeur primaire, possède un certain nombre de règles, comme de faire un choix entre plusieurs codes morphologiques pour une même topographie.
9. Pour le choix entre deux **classifications**, il a été privilégié l'aspect interopérable avec l'existence ou non de tables de correspondance avec les autres classifications, le caractère international de la classification, les recommandations de l'Agence de numérique en santé, les habitudes d'utilisation dans les hôpitaux en France, la granularité sur le sujet traité, le maintien à jour de la classification.
10. Des **conventions** sont parfois nécessaires ; elles sont explicitées et justifiées. Par exemple, une date de début d'évènement tumoral ne correspond pas à une date réelle ;

une maladie métastatique n'apparaît pas comme une maladie aiguë à un instant  $t$  mais correspond à un processus biologique qui peut s'étaler sur plusieurs années. La date de survenue ne peut qu'être qu'une convention. On propose de prendre les conventions internationales quand elles existent.

## OSIRIS RWD : des utilisations multiples

Il existe des utilisations et des finalités diverses avec des conséquences techniques et juridiques selon que les bases de données sont utilisées pour la recherche, pour l'évaluation des soins ou le pilotage de l'activité.

OSIRIS-RWD doit servir les besoins au niveau de chaque établissement, pour les soins, mais aussi permettre des études de recherche multicentriques.

OSIRIS RWD peut permettre la mise en œuvre d'une interopérabilité entre les systèmes d'information hospitaliers (SIH).

OSIRIS RWD peut devenir le socle des entrepôts de données de santé (EDS) hospitaliers qui regroupent des données de différentes sources.

OSIRIS RWD peut servir au niveau local avec un retraitement des données en consolidant des informations de différentes sources pour permettre :

- Le développement d'outils d'aide à la décision, d'aide à la prescription ou au diagnostic
- L'utilisation des données pour une analyse multimodale,
- La mise en application de jumeaux numériques
- Et permettre tout projet impliquant des algorithmes d'intelligence artificielle avec retour au dossier de soins dans une « boucle fermée »

C'est une des raisons qui nécessitent une conversion des données d'OSIRIS RWD au format FHIR.

## Gestion des dates dans le modèle

La gestion des dates est primordiale dans un modèle événementiel qui reconstitue la chronologie de l'histoire de la maladie.

Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY), le champ correspondant sera laissé vide. A titre d'exemple, si l'information du jour n'est pas disponible, elle ne sera pas remplacée par le 15 du mois ou si le mois n'est pas disponible, elle ne sera pas remplacée par la valeur 06.

## Lieu de réalisation des actes médico-techniques et des traitements

Le lieu de prise en charge permet de distinguer si la source d'information est l'établissement de santé ou si l'information provient d'un autre établissement.

Si la prise en charge du cancer est partiellement réalisée en dehors de l'établissement qui alimente la base de données, il est important de recueillir le maximum d'information disponible mais en spécifiant si la prise en charge d'un acte diagnostic ou d'une thérapeutique est fait dans l'établissement (« in ») ou en dehors (« out »).

Les conséquences sont multiples :

- En terme de qualité de données, l'information est toujours meilleure lorsqu'elle provient de l'établissement. Elle est parfois structurée à la source (par exemple pour les actes chirurgicaux) et est plus complète.
- A l'utilisation, certaines études peuvent restreindre l'inclusion aux patients pris en charge dans l'établissement.
- La prise en charge dans l'établissement indique indirectement que des éléments du dossier ou des éléments dérivés des prélèvements sont disponibles (par exemple, si un prélèvement est réalisé dans l'établissement, celui-ci a l'obligation de conserver les échantillons comme les blocs de tumeurs inclus en paraffine).
- Une prise en charge diagnostique ou thérapeutique peut être réalisée en dehors de l'établissement ; elle est alors répertoriée sous la forme d'un compte-rendu conservé dans le dossier électronique (scanné puis interprété par reconnaissance optique des caractères), ou bien l'information est retranscrite dans un compte-rendu médical. Une des problématiques est qu'il n'est pas possible d'assurer l'exhaustivité de l'information.

Si la prise en charge d'un cancer est totalement réalisée en dehors de l'établissement, et avant la première venue du patient dans l'établissement, on peut considérer qu'il s'agit d'un antécédent de cancer. En général, les informations recueillies sont assez succinctes mais cette information mérite d'être recueillie, ne serait-ce que pour repérer les cancers multiples chez un même patient.

## MDS et données complémentaires

Il est prévu de créer une table « observation » pour des données complémentaires non prévues dans le MDS. Cette table pourra contenir des informations cliniques attachées au patient si cela concerne directement la personne : données obtenues dans le cadre d'un examen, d'un interrogatoire ou d'une procédure, sur le mode de vie, de la personne, etc.

Les données sources seront issues d'une paire clé-valeur auquel on ajoute l'identifiant du patient et une date. Pour certaines données, il sera utile de disposer du meilleur rattachement possible de la valeur, au patient, à la tumeur, à l'évolution tumorale, au traitement.

## Evolution du document

Cette version dite 2024.1 est amenée à évoluer.

Issu d'un groupe d'expert dont la liste figure en annexe, il doit être soumis à discussion auprès de différentes institutions ainsi qu'auprès de sociétés savantes.

Il est prévu d'implémenter une première version du modèle dans plusieurs projets de recherche multicentriques afin d'en tester la pertinence, la robustesse et de préciser certains points.

Une traduction en langue anglaise est prévue.

# LE PATIENT

## 1- Identification du patient

### 1.1 Numéro du patient dans la base OSIRIS RWD

Nom de l'item	Numéro du patient dans la base
Numéro de l'item	1.1
Définition	Numéro d'identification unique du patient dans la base OSIRIS RWD
valeur attendue	Un « numéro » pseudonymisé constitué d'une suite de caractères (chiffres)
Règle	<p>Quelle que soit l'utilisation d'OSIRIS RWD, les données identifiantes doivent être pseudonymisées préalablement à leur intégration dans la base de données.</p> <p>La pseudonymisation permet de traiter les données d'individus sans pouvoir identifier ceux-ci de façon directe. Il est toutefois possible de retrouver l'identité de ceux-ci grâce à une table de correspondance.</p> <p>L'opération de pseudonymisation est réversible contrairement à l'anonymisation.</p> <p>Il est de la responsabilité du responsable de traitement d'étudier la meilleure approche pour la mise en œuvre d'une solution de pseudonymisation.</p>
justification du choix de la règle	L'identifiant permanent du patient dans un établissement ne peut être réutilisé comme identifiant.
caractère obligatoire ou optionnel	Obligatoire
Recommandations	Il est recommandé de créer une table de correspondance entre les pseudonymes dans la base de données et les données directement identifiantes.



## 1.2. Identifiant de l'établissement alimentant la base OSIRIS RWD

Nom de l'item	Identifiant de l'établissement alimentant la base OSIRIS RWD
Numéro de l'item	1.2
définition	Numéro FINESS de l'établissement fournisseur des données
Nomenclature utilisée	<p>Le code FINESS (Fichier National des Etablissements Sanitaires et Sociaux) est le répertoire national de référence des personnes morales intervenant dans les domaines sanitaire, médico-social et social.</p> <p>Chaque entité inscrite au répertoire FINESS dispose d'un numéro d'identité unique et définitif composé de 9 caractères, les deux premiers chiffres désignant le département d'implantation.</p> <p>Chaque structure détient 2 numéros FINESS distincts :  Le code FINESS EJ (entité juridique) pour la personne morale titulaire des autorisations d'activité.  Le code FINESS EG (entité géographique) pour les établissements où sont organisés les activités de soins.</p>
valeur attendue	chaîne de caractères
Règle	Le code FINESS utilisé est le code FINESS EG c'est-à-dire celui de l'établissement qui fournit la donnée en d'autres termes, la structure qui alimente la base OSIRIS RWD.
justification du choix de la règle	Historiquement dans OSIRIS, il y avait deux codes FINESS, l'établissement fournisseur de la donnée et l'établissement qui réalise le soin. Il est décidé de ne garder que l'établissement fournisseur de la donnée car le code FINESS de la structure qui effectue le soin n'est pas toujours connu.
caractère obligatoire ou optionnel	obligatoire
Recommandations	Ce code est saisi une seule fois au moment de la création de la base OSIRIS RWD. Ce code FINESS est utile dans les études multicentriques qui consolident les informations venant de plusieurs bases
comparaison avec d'autres standards	<p>Dans le standard OMOP, on trouve dans la table patient, l'ID du fournisseur de données (l'établissement ou même le service) qui sont eux-mêmes enregistrés dans une table « Etablissements de santé ».</p> <p>Dans cette table, les hiérarchies et les professionnels de santé qui fournissent les soins (médecins, infirmières, ...) sont enregistrés dans une table.</p>

### 1.3 Date de naissance

Nom de l'item	Date de naissance
Numéro de l'item	1.3
Définition	Date de naissance de la personne
valeur attendue	YYYY – MM - DD
Règle	La source est le dossier administratif du patient.
caractère obligatoire ou optionnel	Obligatoire
comparaison avec d'autres standards	Dans le modèle OMOP, les informations relatives à une date (année/mois/jour) sont enregistrées dans 3 champs séparés

### 1.4 Sexe du patient

Nom de l'item	Sexe du patient
Numéro de l'item	1.4
Définition	sexe biologique du (de la) patient(e)
valeur attendue	M = sexe masculin F = sexe féminin
Règle	La source est le dossier administratif du patient
justification du choix de la règle	Par convention, les valeurs M ou F sont préférés aux valeurs 1 ou 2.
caractère obligatoire ou optionnel	obligatoire

## 2 - Information et non-opposition du patient pour la ré-utilisation de ses données de santé

### 2.1 Information du patient

Nom de l'item	Information du patient
Numéro de l'item	2.1
définition	Le patient a été informé de la réutilisation de ses données pour la

	recherche.
valeur attendue	Oui/Non
Règle	Cette information doit être tracée et consignée dans le dossier du patient de l'établissement
justification du choix de la règle (exemple)	Les personnes doivent être informées par le responsable de traitement que les données collectées lors de leur prise en charge sont versées au sein de la base de données.
caractère obligatoire ou optionnel	Optionnel
Recommandations	<p>L'établissement a informé le patient de la réutilisation de ses données pour la recherche.</p> <p>Il est de la responsabilité de chaque établissement de tracer cette information et de la mettre à disposition.</p>
Options	<p>Le responsable de traitement peut faire valoir une exception à l'obligation d'information individuelle pour la constitution de la base de données s'il justifie dans son registre d'activité de traitement que la fourniture des informations exigerait des efforts disproportionnés.</p> <p>À ce titre, peuvent notamment être invoqués, au vu de sa situation : le nombre de personnes concernées, l'ancienneté des données, le coût et le temps de la délivrance des informations.</p>

## 2.2 Date d'information du patient

Nom de l'item	Date d'information
Numéro de l'item	2.2
définition	Date d'information de la personne de l'enregistrement de ses données dans la base de données.
Règle	Aucune règle ne sera appliquée. Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
valeur attendue	YYYY - MM - DD
justification du choix de la règle (exemple)	Cf en Annexe Référentiel CNIL – 8. Information des personnes
caractère obligatoire ou optionnel	Optionnel
Recommandations	Cette information n'est pas toujours recueillie dans les hôpitaux.

	<p>Cette information sur la date d'information est optionnelle car secondaire par rapport au fait que le patient soit informé.</p> <p>Il se peut que l'information de la création d'un EDS par l'établissement ne se fasse pas de manière individuelle. Dans ce cas, il s'agit d'une information collective par une campagne de communication globale (livret d'accueil, site web par exemple) et la date de l'information n'est pas forcément pertinente dans ce cas.</p>
--	--

## 2.3 Non-opposition pour l'utilisation des données

Nom de l'item	Non-opposition pour l'utilisation des données
Numéro de l'item	2.3
définition	La personne ne s'oppose à la réutilisation de ses données de santé.
Valeur(s) attendue(s)	Oui/Non
Règle	Par défaut à l'information du patient, on enregistre «Oui» si ce dernier ne manifeste pas son droit d'opposition.
caractère obligatoire ou optionnel	Obligatoire
Recommandations	<p>L'information est souvent présente dans le système d'information des hôpitaux.</p> <p>Il est de la responsabilité de chaque établissement de tracer cette information.</p> <p>Il n'est pas nécessaire d'un consentement spécifique pour les données génétiques ; le consentement pour les données génétiques et les données cliniques sont réunis dans un seul item.</p> <p>Si les données d'un patient sont présentes dans la base de données et que le patient décide de s'opposer à la réutilisation de ces données. C'est de la responsabilité de l'établissement de tracer l'information puis de supprimer les données du patient concerné.</p>
Options	Si la personne s'oppose à la ré-utilisation de ses données, le patient peut ne pas figurer dans l'EDS et l'alimentation des données de cette personne sont bloquées à la source de la base.

## 2.4 Date d'opposition pour l'utilisation de ses données

Nom de l'item	Date d'opposition pour la réutilisation de ses données
Numéro de l'item	2.4
définition	Date de signature de la demande d'opposition du (de la) patient(e) pour l'utilisation de ses données cliniques et génétiques
valeur attendue	YYYY - MM - DD
Règle	Aucune règle ne sera appliquée. Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
justification du choix de la règle (exemple)	Cf en Annexe Référentiel CNIL – 9. Droits des personnes
caractère obligatoire ou optionnel	Optionnel
Options	Se pose la question de l'utilité de cet item puisqu'on peut penser que si le patient s'est opposé à son enregistrement dans l'EDS il ne doit plus y figurer. Pour autant, le patient peut accepter puis signifier son refus. Ainsi, toute étude réalisée avec son consentement doit être tracée, ce qui implique de conserver la date d'opposition.

## 3 - Facteurs de risque

### 3.1 Facteurs de risque

Nom de l'item	Facteurs de risque
Numéro de l'item	3.1
Définition	Facteurs de risque externes
Nomenclature utilisée	Codification propre à OSIRIS RWD Les facteurs de risque sont exprimés sous forme d'information binaire, patients présentant ou non tel facteur de risque
valeur attendue	Oui/Non/Inconnu (par variable)  Il est proposé une liste fermée : <ul style="list-style-type: none"><li>- Tabagisme (dont tabagisme passif)</li><li>- Consommation d'alcool</li><li>- Alimentation déséquilibrée</li><li>- Surpoids et obésité</li><li>- Activité physique insuffisante</li></ul>

	<ul style="list-style-type: none"> <li>- Utilisation d'hormones exogènes</li> <li>- Durée insuffisante de l'allaitement infantile</li> <li>- Infections (ex : hépatite B et C, papillomavirus, H. Pylori)</li> <li>- Radiations ionisantes (radon, exposition diagnostic)</li> <li>- Pollution atmosphérique, particules fines</li> <li>- Rayonnement solaire (UV)</li> <li>- Expositions professionnelles (ex : amiante)</li> <li>- Exposition à certaines substances chimiques (ex : arsenic, benzène).</li> </ul>
Règle	La collecte de l'information se fait par NLP à partir des dossiers textuels ou en récupérant une information structurée.
justification du choix de la règle (exemple)	La liste proposée est réalisée avec les propositions des facteurs de risque de l'INCA basée sur la liste du CIRC.
caractère obligatoire ou optionnel	Optionnel
Recommandations	<p>Le cancer est une maladie multifactorielle et il existe de nombreux facteurs de risque qui peuvent contribuer au développement de différents types de cancer. Certains sont identifiés et démontrés, de nombreux facteurs sont suspectés.</p> <p>Les facteurs de risque reconnus sont le tabagisme, la consommation d'alcool, un surpoids, l'alimentation trop riche, une exposition à certaines substances toxiques (amiante), aux rayonnements ionisants ou certaines infections virales et bactériennes (papillomavirus, virus de l'hépatite B et C.</p> <p>Certains des facteurs de risque seront aisés à structurer de façon automatique comme par exemple le « tabagisme » ou une « consommation d'alcool » car l'information existe très souvent sous forme textuelle ou structurée dans les dossiers (première consultation ou consultation d'anesthésie par exemple).</p> <p>D'autres facteurs de risque seront certainement moins évidents à récupérer comme « alimentation déséquilibrée » ou « activité physique insuffisante ».</p>

## 4 - Statut vital du patient

Les informations sur le statut vital du patient sont issues de plusieurs sources :

- en l'absence d'information, le patient est considéré comme vivant (à condition de vérifier qu'il n'est pas décédé par exemple dans le fichier INSEE) ; la source d'information utilisée est nécessaire
- si le patient s'est présenté dans l'établissement à une date donnée, il est vivant
- si une information est disponible dans le dossier par exemple par le biais d'un courrier, l'information peut être récupérée.

La recommandation est de conserver trois dates distinctes : la date de dernière visite, la date de dernière nouvelle et la date de décès.

### 4.1 Statut vital

Nom de l'item	Statut vital
Numéro de l'item	4.1
définition	Statut vital de la personne
valeur attendue	VIVANT / DECEDE
Règle	<p>Le système récupère l'information dans le dossier administratif (ou dans le dossier médical)</p> <p>Par défaut le patient est mis en statut « vivant » jusqu'à ce qu'on remonte l'information qu'il est décédé soit du dossier administratif de l'établissement, du dossier médical ou via les données de l'INSEE</p>
justification du choix de la règle (exemple)	NA
caractère obligatoire ou optionnel	Obligatoire
Recommandations	Pour une étude particulière, il est recommandé, si le patient n'est pas déclaré « décédé », d'interroger le fichiers des personnes décédées de l'INSEE (Lien : <a href="https://www.insee.fr/fr/information/4190491">https://www.insee.fr/fr/information/4190491</a> ) pour savoir si le patient n'est pas inscrit dans le fichier INSEE comme décédé.

### 4.2 Source de l'information « Statut vital »

Nom de l'item	Source de l'information « Statut vital » récupérée auprès de l'INSEE
Numéro de l'item	4 .2
définition	Information concernant le statut vital de la personne récupérée depuis le fichier INSEE

valeur attendue	Oui / Non
Règle	Il est nécessaire de consulter deux sources pour renseigner cette information. La première source est le dossier médical du patient (ou le dossier administratif) puis le fichier de l'INSEE.
caractère obligatoire ou optionnel	optionnel
Recommandations	<p>Les algorithmes de recherche pour un appariement un patient à l'établissement et la base INSEE sont différents entre les établissements.</p> <p>La qualité de l'information issue de l'INSEE dépend directement de l'algorithme et des critères de filtre.</p> <p>La recommandation pourrait être d'utiliser les mêmes méthodes avec un intérêt réel pour les études multicentriques.</p>

### 4.3 Date de dernière venue

Nom de l'item	Date de dernière venue
Numéro de l'item	4.3
définition	Date de dernière venue physique de la personne dans l'établissement
valeur attendue	YYYY - MM - DD
Règle	<p>L'établissement s'entend par l'établissement qui alimente la base OSIRIS RWD.</p> <p>Il ne s'agit pas d'une date de rendez-vous mais bien de la venue physique dans l'établissement.</p> <p>Cette date est alimentée en automatique à partir du dossier du patient dans l'établissement.</p> <p>L'établissement choisit la source la plus fiable ; la source peut être celle qui déclenche une facturation ou celle qui définit qu'une venue physique a bien été réalisée.</p>
justification du choix de la règle	Cette date est réputée fiable car récupérée à chaque venue avec des objectifs de facturation
caractère obligatoire ou optionnel	Obligatoire
Recommandations	La date de dernière visite dans l'établissement est une donnée toujours disponible et fiable soit parce que il y a eu une consultation, une hospitalisation ou un passage en hôpital de jour.
Option non retenue	Il n'est pas retenu le principe qu'une date de document soit considéré comme une date de dernière venue.



## 4.4 Date des dernières nouvelles

Nom de l'item	Date des dernières nouvelles
Numéro de l'item	4.2
définition	Date des dernières nouvelles à travers les courriers que l'établissement peut recevoir après la dernière visite du patient. Il ne s'agit pas de la date du document mais d'une date extraite du contenu des courriers.
valeur attendue	YYYY - MM - DD
Règle	La collecte de cette information par OCR (reconnaissance optique de caractères) sur les CR extérieurs reçus par l'établissement et que l'établissement aura scanné dans le dossier patient.  Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
justification du choix de la règle (exemple)	Il ne s'agit pas de la date du document mais d'une date extraite du contenu des courriers, détectée par NLP
caractère obligatoire ou optionnel	Optionnel
Recommandations	L'item est distingué de la date de dernière visite physique dans l'établissement.
Options	Si on ne détecte pas la date dans le document, une option consiste à utiliser la date de scannage du document
comparaison avec d'autres standards	OMOP - Pour les séjours en hospitalisation, la date de fin correspond généralement à la date de sortie. Si une personne est toujours hospitalisée à l'établissement au moment de l'extraction des données et n'a pas de visit_end_date, il est proposé visit_end_date sur la date de l'extraction des données.

## 4.5 Date de décès

Nom de l'item	Date de décès
Numéro de l'item	4.5
définition	Date à laquelle la personne est décédée.
valeur attendue	YYYY - MM - DD

Règle	<p>En théorie, l'information est disponible dans le dossier médical du patient et/ou dans son dossier administratif</p> <p>En pratique, l'information peut être présente dans le dossier médical, le dossier administratif avec possiblement une incohérence</p> <p>Cette information est parfois absente alors que le patient est décédé.</p>
caractère obligatoire ou optionnel	Optionnel
Recommandations	<p>La très grande majorité des patients décèdent à l'extérieur de l'établissement.</p> <p>L'information la plus fiable est celle qui figure dans les bases de l'INSEE et pour chaque étude, il serait nécessaire d'interroger la base INSEE</p> <p>La recommandation est de noter la source au niveau de l'item 4.2.</p> <p>On distingue deux sources d'alimentation de la base de données OSIRIS-RWD concernant la date de décès, la première c'est l'établissement lui-même par ses dossiers administratifs et une seconde source par l'INSEE.</p> <p>Si l'information vient de l'INSEE, l'information est mise à jour dans le champ correspondant. Dans ce cas, la date de l'INSEE prévaut et remplace la date de décès de l'établissement.</p>
comparaison avec d'autres standards	OMOP- si le mois de la date du décès n'est pas connu ou n'est pas autorisé alors le mois de décembre de l'année est utilisé par défaut ainsi que le dernier jour du mois.

## 5 - Antécédents familiaux de cancer

### 5.1 Topographie de l'antécédent familial de cancer

Nom de l'item	Topographie de l'antécédent familial de cancer
Numéro de l'item	5.1
définition	Localisation de l'antécédent carcinologique familial selon la CIM-10
Nomenclature utilisée	CIM-10
valeur attendue	chaîne de caractères (ex:C18.4)
caractère obligatoire ou optionnel	Optionnel

## 5.2 Morphologie de l'antécédent familial de cancer

Nom de l'item	Morphologie de l'antécédent familial de cancer
Numéro de l'item	5.2
définition	Description morphologique de l'antécédent carcinologique familiale selon la Classification Internationale des Maladies pour l'Oncologie (CIM-O- version 3).
Nomenclature utilisée	CIM-O-3 Morpho
valeur attendue	chaîne de caractères (ex : 8144/3)
caractère obligatoire ou optionnel	Optionnel

## 5.3 Lien de parenté avec le patient

Nom de l'item	Lien de parenté
Numéro de l'item	5.3
définition	Type de relation entre la personne et l'antécédent familial
Nomenclature utilisée	Liste fermée
valeur attendue	Parenté au 1er degré : <ul style="list-style-type: none"><li>- Mère</li><li>- Père</li><li>- Fils</li><li>- Fille</li></ul> Parenté au 2ème degré : <ul style="list-style-type: none"><li>- Grand-père</li><li>- Grand-mère</li><li>- Petit-fils</li><li>- Petite-fille</li><li>- Frère</li><li>- Sœur</li></ul>
caractère obligatoire ou optionnel	Optionnel
Recommandations	La recommandation est de structurer toutes les valeurs retrouvées dans la liste du 1 <sup>er</sup> degré et du 2 <sup>nd</sup> degré de parenté du code civil.

## 6 - Antécédent personnel de cancer

Les antécédents personnels de cancer sont comptabilisés comme des occurrences de cancer (Cf. chapitre 8 - Le cancer).

## 7 - Antécédent médical, comorbidité et événement indésirable

### Définitions

Un **antécédent médical** est la survenue d'une maladie ou d'un symptôme, aiguë ou chronique, avant une date fixée, par exemple la date de diagnostic d'un cancer. Le patient peut présenter, avant la survenue de la maladie cancéreuse, une maladie qui peut influencer la décision thérapeutique.

Une **comorbidité** est la survenue d'une maladie ou d'un symptôme, aiguë ou chronique, survenant au cours de la prise en charge d'une pathologie, en l'occurrence la maladie cancéreuse. Des événements pathologiques, les comorbidités, sont, soit favorisés par la maladie cancéreuse, soit seraient apparus de toute façon. A l'évidence, il est difficile d'affirmer ce qu'il serait advenu indépendamment du cancer.

Un **événement indésirable** est la conséquence, souvent prévisible, d'un traitement survenant en plus de son effet principal. Il existe des effets secondaires immédiats et des effets secondaires tardifs. Un effet secondaire peut persister longtemps après l'arrêt des traitements et, parfois, devenir définitif.

### Exemples

Trois exemples peuvent illustrer cette problématique ; la maladie thrombo-embolique veineuse (MTEV), l'insuffisance rénale, des troubles digestifs plus ou moins sévères.

#### *La maladie thrombo-embolique veineuse (MTEV)*

Une MTEV, embolie pulmonaire ou thrombose veineuse profonde, est une pathologie fréquente en population générale. Si elle survient avant la découverte du cancer, on la définira comme un antécédent médical qui pourra influencer la thérapeutique du cancer ; certaines hormonothérapies augmentent le risque de MTEV et on pourra privilégier d'autres classes thérapeutiques dont le risque est moindre.

Si la MTEV survient au cours de la maladie cancéreuse, on parlera de comorbidité, d'autant que l'on sait que pour certains cancers, le risque est multiplié par un facteur 2 à 3. Ceci est d'autant plus important que la MTEV est la deuxième cause de mortalité des patients atteints de cancer.

#### *L'insuffisance rénale chronique (IRC)*

L'insuffisance rénale chronique (IRC) est une maladie fréquente qui toucherait, aux **Etats-Unis**, 15 % des adultes. **En France**, une étude chez plus de 9000 patients âgés de 65 ans ou plus, estimait la fréquence d'IRC à 13,7% (DFG inférieur à 60 ml/min/1,73m<sup>2</sup>). Or, certains médicaments du cancer favorisent la survenue d'une IRC et peuvent être considérés comme un effet indésirable dont la gravité est variable, allant d'une simple baisse de la DFG à une insuffisance rénale aiguë. En fonction de la date de survenue, on définira l'IRC comme un antécédent ou comme un effet indésirable potentiel. De plus, il ne peut être exclu qu'il s'agit d'une comorbidité qui serait apparu de toute façon, soit avec une cause connue (liée à une prise médicamenteuse comme des AINS), soit d'étiologie inconnue.

## Les troubles digestifs

Les troubles digestifs, vomissements et diarrhées, sont des complications fréquentes des traitements de chimiothérapie ou de radiothérapie. Ils peuvent aussi survenir lors à l'occasion d'une intoxication alimentaire chez tout un chacun. Si ces symptômes surviennent dans les suites immédiates d'un traitement, on parlera volontiers d'effets indésirables. S'ils surviennent à distance, c'est plus probablement une cause fortuite, indépendante de tout traitement.

## Application au modèle OSIRIS RWD

Ces trois exemples montrent la difficulté (voire l'impossibilité) d'imputer des comorbidités ou des effets indésirables aux traitements du cancer, surtout lors d'associations thérapeutiques (bithérapie ou association radiothérapie-chimiothérapie).

La convention adoptée est de recueillir l'information brute appelée **évènement morbide** sans préjuger du fait qu'il s'agit d'un antécédent médical, une comorbidité ou un effet indésirable. La date de survenue, de découverte ou de recueil dans le dossier médical sera notée, à charge au moment de l'analyse d'interpréter les résultats comme un antécédent, une comorbidité ou un possible effet indésirable.

## Evènements morbides et choix du référentiel

Les antécédents médicaux ou les comorbidités sont notés comme des maladies. La classification utilisée sera la CIM-10

Les effets indésirables sont le plus souvent des symptômes. La classification utilisée est une partition de la CTCAE.

## Gestion des dates

Les antécédents médicaux ou les comorbidités sont notés dans le dossier médical. La date de début est parfois indiquée en clair mais souvent de façon imprécise.

Parfois, l'information est notée clairement comme un antécédent, par exemple, lors de la première venue ou lors de la première consultation.

L'information peut apparaître à n'importe quel moment, dans tout type de compte rendu médical. Si la date est précisée dans le compte-rendu, on utilisera des techniques de traitement du langage qui permettent d'attribuer la bonne date à l'évènement. S'il n'est pas possible d'affecter une date dans le corps du compte-rendu, on propose d'affecter à l'évènement, la date de validation du compte-rendu.

Pour les effets indésirables, la précision quant à la date d'apparition est inconstante. Si cela est précisée on note la date de survenue de l'évènement ; dans le cas contraire, c'est la date du compte-rendu qui fera office de date de survenue.

## Données unitaires ou données répétitives

Une donnée concernant une maladie chronique ne peut être répétitive, par définition comme par exemple l'insuffisance rénale chronique. Par contre, une comorbidité aiguë peut récidiver : il est possible de présenter plusieurs épisodes d'embolie pulmonaire. A l'évidence, les effets indésirables sont très souvent répétés dans le temps, surtout s'ils sont liés à une thérapeutique médicale qui se reproduit comme les plusieurs cycles de chimiothérapie.

Par convention, on propose que les premiers évènements morbides codés en CIM10 soient notés et que tous les évènements morbides à type de symptômes codés en CTCAE sont répétitifs.

## 7.1 Antécédent médical personnel (hors cancer) / comorbidité

Nom de l'item	Antécédent médical personnel (hors cancer) / comorbidité
Numéro de l'item	7.1
définition	Pathologie hors cancer
Nomenclature utilisée	CIM 10
valeur attendue	Chaine de caractères
Règle	La règle est de structurer dans cet item les pathologies associées (comorbidités, antécédents médicaux).  C'est la date de début qui permettra de qualifier et de séparer un antécédent d'une comorbidité.
justification du choix de la règle (exemple)	Cet item permet de distinguer les pathologies associées survenues avant le diagnostic de cancer (antécédent médical du cancer) des pathologies associées survenues pendant la maladie cancéreuse (comorbidités).
caractère obligatoire ou optionnel	Optionnel

## 7.2 Evénement indésirable

Nom de l'item	Evénement indésirable
Numéro de l'item	7.2
définition	Evènement indésirable repéré dans le dossier médical
Nomenclature utilisée	CTCAE
valeur attendue	Chaine de caractères
Règle	Evènement indésirable identifié après l'administration du dernier traitement ou intervention médicale.
justification du choix de la règle (exemple)	Cet item permet de distinguer les pathologies associées survenues avant le diagnostic de cancer (antécédent médical du cancer) des pathologies associées survenues pendant la maladie cancéreuse (comorbidités).

caractère obligatoire ou optionnel	Optionnel
Recommandations	La classification CTCAE permet de coder les événements indésirables mais sa précision est mal adaptée aux données de la vie réelle. Une partition sera proposée avec un nombre limité d'item.

### 7.3 Date de survenue de l'évènement morbide

Nom de l'item	Date de survenue de l'évènement morbide
Numéro de l'item	7.3
Définition	Date à laquelle le symptôme ou la maladie sont notés pour la première fois
valeur attendue	YYYY - MM - DD
Règle	La date de survenue est la date notée dans le dossier  Aucune règle ne sera appliquée. Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
caractère obligatoire ou optionnel	Optionnel

# LE CANCER

## 8 – La tumeur primitive

Un cancer est défini par :

- Une occurrence (en raison des cancers multiples)
- Un lieu de prise en charge
- Une date de diagnostic (anatomo-pathologique ou cytologique)
- Une localisation tumorale (ou code topographique)
- Une anatomo-pathologie (ou morphologie)
- Une latéralité
- Un bilan d'extension initial : TNM et/ou stade
- La taille de la tumeur primaire
- Un grade histologique
- Des biomarqueurs prédictifs, pronostiques ou théranostiques (voir le chapitre biomarqueurs)
- Une date de début de traitement (voir le chapitre traitements)

*Exemple : Patiente sans antécédent de cancer (occurrence 1), présentant un adénocarcinome galactophorique infiltrant du sein droit mesurant 18 mm avec une biopsie réalisée le 1/02/2022 ; tumeur classé T2N0M0 de stade IIA, de grade SBR II, RO + RP – HER2 - dont le traitement a débuté le 1/03/2022.*

Pour les dates de début de la maladie cancéreuse, deux dates sont conservées :

- La date de diagnostic du cancer, le plus souvent histologique, cytologique, plus rarement sur des tests cytogénétiques ou un examen d'imagerie.
- La date de début de premier traitement du cancer (figure dans le chapitre « Traitement »).

La date dite de découverte du cancer (par exemple basée sur les premiers symptômes, un examen clinique, une mammographie de dépistage, un scanner thoracique pour un cancer du poumon, une IRM pour le cerveau) n'est pas conservée dans le MDS car sujet à trop de disparité.

Il est recommandé de structurer à part les tumeurs bénignes (codes D10-D36) et de ne pas les intégrer dans le modèle.

### 8.1 Occurrence du cancer

Nom de l'item	Occurrence du cancer (numéro d'ordre)
Numéro de l'item	8.1
Définition	Ordre d'apparition du cancer pour le patient, quel que soit le lieu de prise en charge
règle	Le numéro du cancer correspond à l'ordre des cancers intervenus



	<p>dans la vie de la personne.</p> <p>Pour les cancers pris en charge dans l'établissement, la date de diagnostic permet de discerner l'ordre.</p> <p>Pour les antécédents de cancer, l'année présente dans le dossier médical est prise en compte.</p>
valeur attendue/ vocabulaire attendu	Entier de 1 à n
justification du choix de la règle (exemple)	<p>Le numéro du cancer 1, 2, 3, etc. correspond à l'ordre des cancers intervenus dans la vie des patients quel que soit le lieu de prise en charge.</p> <p><i>Exemple : Une patiente vient dans l'établissement A pour un cancer du sein gauche. Elle a eu auparavant un cancer du sein droit il y a 10 ans traité dans l'établissement B.</i></p> <p><i>Le cancer sein droit doit être enregistré en numéro 1.</i></p> <p><i>Le cancer du sein gauche doit être enregistré en numéro 2.</i></p>
Obligatoire/optionnel	Obligatoire (calculé)
recommandation	Le numéro du cancer (ou occurrence) est calculé automatiquement à partir des informations sur la date du diagnostic, le code topographique et la latéralité
comparaison avec d'autres standards	Dans OMOP, c'est l'algorithme d'OMOP qui remplit le champ episode_number dans la table EPISODE au moment du remplissage du champ diagnostic.

## 8.2 Prise en charge dans l'établissement

Nom de l'item	Prise en charge dans l'établissement
Numéro de l'item	8.2
définition	Prise en charge totale ou partielle (au moins un acte) du patient dans l'établissement pour l'occurrence de cancer
règle	L'établissement dans lequel une prise en charge totale ou partielle (au moins un acte) a lieu est considéré comme un lieu de prise en charge.
valeur attendue/ vocabulaire attendu	OUI / NON
justification du choix de la règle (exemple)	<p><i>Un patient a été pris en charge intégralement pour un cancer en dehors de l'établissement sans jamais venir pour aucun acte diagnostic, thérapeutique ou une surveillance, il est considéré que la prise en charge se fait en dehors de l'établissement. S'il vient ensuite pour une 2<sup>ème</sup> tumeur et est pris en charge dans l'établissement, il est considéré comme pris en charge pour la 2<sup>ème</sup> tumeur, même si cette prise en charge est partielle (acte diagnostic ou thérapeutique, deuxième avis).</i></p>

	<p><i>Un patient vient dans l'établissement A pour un cancer du poumon et est traité par chirurgie puis radiothérapie. Il est considéré comme pris en charge dans l'établissement.</i></p> <p><i>Un patient est traité par chimiothérapie dans l'établissement A ; la radiothérapie est réalisée dans l'établissement B. Il est considéré comme pris en charge dans l'établissement A et dans l'établissement B.</i></p> <p><i>Un patient a été traité pour un cancer ORL dans l'établissement A sans jamais venir dans l'établissement B. Secondairement, il est ensuite pris en charge dans l'établissement B ; le cancer ORL est considéré comme pris en charge en dehors de l'établissement ; le cancer d'occurrence 2 est considéré comme pris en charge dans l'établissement B.</i></p>
option	Pour les antécédents de cancer, la modalité « Non » est retenue.
comparaison avec d'autres standards	OMOP onco, c'est l'algorithme d'OMOP qui remplit le champ episode_number dans la table EPISODE au moment du remplissage du champ diagnostic.

### 8.3 Date de diagnostic du cancer

Nom de l'item	Date de diagnostic du cancer
Numéro de l'item	8.3
Définition	<p>Date de l'examen ou de l'acte (qui inclut l'acte chirurgical si c'est l'acte diagnostic) qui permet le diagnostic de certitude du cancer.</p> <p>La date de diagnostic correspond au premier diagnostic du cancer en sachant que le cancer est éventuellement métastatique d'emblée.</p>
valeur attendue	YYYY - MM - DD
Règle	<p>On priorise pour le choix de la date de diagnostic, la date de l'examen selon l'ordre suivant :</p> <p>(1) la date d'histologie (date de prélèvement ou la date de la chirurgie si l'examen histologique a été pratiquée à la suite d'un acte chirurgical)</p> <p>(2) la date de la cytologie</p> <p>(3) la date des prélèvements de cytogénétiques</p> <p>(4) la date de l'imagerie diagnostique</p> <p>Dans tous les cas, c'est la date de prélèvement ou de l'examen qui prévaut et non la date du résultat de l'examen.</p> <p>Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.</p>
caractère obligatoire ou optionnel	Obligatoire

	Si antécédent, année obligatoire
Remarque	<p>Cas particulier des cancers synchrones :</p> <p>Il s'agit d'un des cas particuliers parmi les plus difficiles.</p> <p>Il s'agit de situations assez rares (moins de 1 % des situations), et difficilement exploitables en recherche.</p> <p>Il est nécessaire d'assurer leur repérage et de bien les définir, ne serait-ce que pour les exclure dans certaines recherches.</p>

## 8.4 Méthode de diagnostic du cancer

Nom de l'item	Méthode de diagnostic du cancer
Numéro de l'item	8.4
Définition	Méthode d'investigation et de confirmation du diagnostic de cancer
Règle	Dès qu'une investigation avec un ordre plus élevé (cf. infra) est réalisée, cette modalité prime sur les codes inférieurs.
Nomenclature utilisée	Basée sur les règles européennes de codage dans les registres. ( <a href="https://encr.eu/sites/default/files/Recommendations/ENCR-Recommendation-standard-dataset_Mar2023.pdf">https://encr.eu/sites/default/files/Recommendations/ENCR-Recommendation-standard-dataset_Mar2023.pdf</a> )
Valeurs attendues	<p>1 clinique</p> <p>2 investigation clinique (y compris imagerie)</p> <p>3 biomarqueurs spécifiques</p> <p>4 pour cytologie</p> <p>5 pour histologie</p> <p>6 pour les tests cytogénétiques</p> <p>0 pour inconnu</p>
caractère obligatoire ou optionnel	Optionnel
Comparaison avec d'autres standards	<p>Le NCI propose dans le manuel « SEER Program Coding and Staging Manual 2024 » une classification différente pour les tumeurs solides et les tumeurs hématologiques (page 111)</p> <p><a href="https://seer.cancer.gov/manuals/2024/SPCSM_2024_MainDoc.pdf">https://seer.cancer.gov/manuals/2024/SPCSM_2024_MainDoc.pdf</a></p>

## 8.5 Code topographique

Nom de l'item	Code topographique du cancer
Numéro de l'item	8.5
Définition	Localisation de la tumeur primitive.

Nomenclature utilisée	CIM-10
valeur attendue	chaîne de caractères (ex : C18.4)
Règle	<p>Le code topographique décrit le site anatomique d'origine de la tumeur.</p> <p>Le code a toujours un préfixe de « C », suivi d'un nombre à trois chiffres qui indique le site (deux chiffres) et le sous-site (un chiffre), séparés par un point décimal.</p> <p>Exemple : C18.4 : le C18 indique que le site est le côlon et le 4 indique que le sous-site est le côlon transverse.</p>
caractère obligatoire ou optionnel	Obligatoire
Options	<p>Pour les lymphomes ganglionnaires, le code approprié est C77.8 (Ganglion lymphatique de territoires multiples) ou C77.9 pour “ ganglion lymphatique SAI ”.</p> <p>Les lymphomes extra-ganglionnaires doivent être codés en fonction de leur localisation d'origine. Si le diagnostic n'indique pas de localisation, et si le lymphome est extra-ganglionnaire, utiliser le code C80.9 (localisation primitive inconnue).</p> <p>Toutes les leucémies doivent être codées en C42.1 (moelle osseuse).</p> <p>Les autres tumeurs (classés en D dans la CIM10) sont recueillies dans le modèle OSIRIS RWD. D37-D48 Tumeurs à évolution imprévisible ou inconnue D00-D09 Tumeurs in situ</p>
Recommandations	<p>La classification topographique de la CIM-O-3 couvre 99 % des problèmes y compris pour</p> <ul style="list-style-type: none"> <li>- Les leucémies qui sont classées en Moelle osseuse</li> <li>- Les sarcomes,</li> <li>- Les lymphomes</li> </ul>
comparaison avec d'autres standards	<p>mCODE : les codes décrivant le ou les emplacements du cancer primaire ou secondaire = ensemble de valeurs qui comprennent tous les codes de la hiérarchie de structure de corps SNOMED CT (codes descendant de 123037004 'Body Structure').</p> <p>La topographie peut être codé à l'aide des codes topographiques de la CIM-O-3.</p> <p>Seuls SNOMED CT et ICD-O-3 sont considérés comme conformes.</p> <p><a href="https://bioportal.bioontology.org/ontologies/SNOMEDCT/?p=classes&amp;conceptid=http%3A%2F%2Fpurl.bioontology.org%2Fontology%2FSNOMEDCT%2F123037004">https://bioportal.bioontology.org/ontologies/SNOMEDCT/?p=classes&amp;conceptid=http%3A%2F%2Fpurl.bioontology.org%2Fontology%2FSNOMEDCT%2F123037004</a></p>

## 8.6 Code histologique / morphologique

Nom de l'item	Code histologique / morphologique
Numéro de l'item	8.6
définition	Description morphologique du cancer
Nomenclature utilisée	CIM-O-3
valeur attendue	chaîne de caractères (ex : 8144/3)
Règle	<p>Les règles utilisées sont celles de la CIM-O-3.</p> <p>Pour une même entité tumorale, on différencie les codes comportementaux (le 5ème digit) et des codes morphologiques (les 4 premiers digits).</p> <p>Si on note plusieurs codes anapath pour une même tumeur, on conserve le code le plus élevé (comme proposé par la classification CIM-O-3).</p> <p>Les registres n'utilisent qu'un seul code morphologique (même si la tumeur est composite ou multiple dans un même site primitif)</p> <p>En cas de tumeurs composites ou multiples (des règles s'imposent car on n'utilise qu'un seul code morphologique)  Ex : si "/X" différents alors on prend toujours le "/X" le plus élevé (cf RULE F de l'ICDO3)  Ex : si "XXXX" différents avec le même "/X" alors on prend le XXXX le plus élevé (cf RULE K de l'ICDO3)</p> <p>Exemples :</p> <ul style="list-style-type: none"> <li>- on conserve toujours un code /3 par rapport à un /2 ou un /1.</li> <li>- s'il y a plusieurs morphologies différentes, on conserve les 4 premiers digits les plus élevés.</li> </ul> <p>Pour les cancers primitifs, le suffixe comportemental de la CIM-O-3 doit être /1, /2 ou /3. Pour les cancers secondaires, le suffixe comportemental de la CIM-O-3 doit être /6.</p> <p>L'application de ces règles permettent de ne conserver qu'un seul code histologique (information uni-variée).</p>
caractère obligatoire ou optionnel	Obligatoire
Options	Si l'établissement utilise la codification ADICAP, il existe un transcodage en code ICD-O-3 dans la mesure où la classification ADICAP n'est pas maintenu ; elle n'est donc pas recommandée.
comparaison avec d'autres standards	mCODE : les codes morphologiques de la CIM-O-3 sont référencés pour représenter la structure, la disposition et les caractéristiques

	comportementales des néoplasmes malins et des cellules cancéreuses. Critères d'inclusion : tumeurs in situ et tumeurs malignes. Critères d'exclusion : tumeurs bénignes et tumeurs de comportement indéterminé.
--	---

## 8.7 Latéralité de la tumeur primaire

Nom de l'item	Latéralité
Numéro de l'item	8.7
définition	Latéralité de la tumeur primitive
Nomenclature utilisée	Nomenclature propre à OSIRIS
valeur attendue	Bilatéral Droit Gauche Médian Non applicable Non connu
caractère obligatoire ou optionnel	Obligatoire
Recommandations	<p>Pour certain cancer comme le cancer du sein, ce champ doit être rempli. Deux options possibles en cas de cancer du sein bilatéral d'emblée. Soit on code « cancer bilatéral », soit on code « deux cancers », un « droit » et un « gauche ». La recommandation est de coder en cancer du sein bilatéral si les deux cancers sont diagnostiqués avant tout traitement.</p> <p>Dans le cas où les cancers sont successifs avec un traitement entre les deux, on crée un cancer d'un côté, puis un deuxième cancer contro-latéral, soit deux cancers</p> <p>L'information peut être « non connu » et pour d'autre « non applicable » comme par exemple les lymphomes.</p>
comparaison avec d'autres standards	<p>Le NCI pour le programme SEER codifie d'une façon différente pour la latéralité et donne des recommandations pour les cancers synchrones</p> <p><a href="https://seer.cancer.gov/manuals/2024/SPCSM_2024_MainDoc.pdf">https://seer.cancer.gov/manuals/2024/SPCSM_2024_MainDoc.pdf</a></p>

## 8.8 Taille de la tumeur primitive au diagnostic

Nom de l'item	Taille de la tumeur primitive au diagnostic
Numéro de l'item	8.8
définition	Permet de décrire la taille de la tumeur primaire (en mm).
valeur attendue	Entier
Règle	<p>La taille de la tumeur primitive (la valeur du plus grand axe) est mesurée par les cliniciens, les radiologues, les pathologistes et l'information est consignée dans le dossier médical.</p> <p>L'unité de la valeur est le millimètre</p> <p>Les méthodes sont de 3 types : clinique, anatomo-pathologique et imagerie.</p>
caractère obligatoire ou optionnel	Optionnel
Recommandations	On ne conserve qu'une seule mesure à savoir la plus grande dimension (le plus grand axe)
Option	Pour les tumeurs multifocales, notamment des cancers du sein, il est proposé de retenir la plus grande taille.
comparaison avec d'autres standards	<ul style="list-style-type: none"> <li>- Variable « date » (YYYY - MM - DD, Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.)</li> <li>- Variable « la dimension la plus longue de la tumeur » en mm</li> <li>- Variable « Méthode de mesure » codée par la Snomed <ul style="list-style-type: none"> <li>▪ Examen physique Snomed CT 5880005</li> <li>▪ Examen pathologique <ul style="list-style-type: none"> <li>• macroscopique (SCT 787377000)</li> <li>• microscopique (SCT 104157003)</li> </ul> </li> <li>▪ Examen d'imagerie <ul style="list-style-type: none"> <li>• Radiologie (SCT 363680008)</li> <li>• Echographie (SCT 16310003)</li> <li>• IRM (SCT 113091000)</li> <li>• tomographie (CAT) (SCT 77477000)</li> <li>• PET scan (SCT 82918005)</li> <li>• Scintigraphie (ex : thyroïde)</li> <li>• Imaging (SCT 363679005)</li> </ul> </li> </ul> </li> </ul>

## 8.9 cotation du T

Nom de l'item	cotation du T
Numéro de l'item	8.9
définition	la cotation du T (qui représente la taille de la tumeur) dans la classification TNM.
Nomenclature utilisée	Notre recommandation est d'utiliser un fichier qui liste l'ensemble des codes de la classification de l'UICC TNM (version en fonction de la date du CR médical)
valeur attendue	chaîne de caractères (liste fermée) Liste de codes de la classification TNM (présente et passées)
Règle	La valeur du T est une valeur unique codée au moment du diagnostic de la tumeur (pour cT et pT) et peut différer selon le type de classification. <ul style="list-style-type: none"> <li>- un seul cT coté par le clinicien avant tout traitement,</li> <li>- un seul pT coté par le pathologiste</li> <li>- un seul yT après traitement néo-adjuvant.</li> </ul>
caractère obligatoire ou optionnel	Optionnel
Recommandations	Lorsque, pour une même tumeur, deux stades de T sont notés dans le dossier médical, on prend la valeur la plus élevée.
comparaison avec d'autres standards	Dans mCODE la cotation du T est dans l'item « TNM Primary Tumor Category Profile » <a href="#">LOINC 21905-5 Primary tumor.clinical [Class] Cancer</a>

## 8.10 cotation du N

Nom de l'item	cotation du N
Numéro de l'item	8.10
définition	la cotation du N (qui représente l'atteinte éventuelle de ganglions lymphatiques régionaux) dans la classification TNM.
Nomenclature utilisée	la classification TNM (diffère selon la version) Notre recommandation est d'utiliser un fichier qui liste l'ensemble des codes de la classification de l'UICC TNM (version en fonction de la date du CR médical)
valeur attendue	chaîne de caractères (liste fermée) Liste de codes de la classification TNM (présente et passées)
Règle	Même règle que pour le T
caractère obligatoire ou optionnel	Optionnel



comparaison avec d'autres standards	Dans mCODE la cotation du N est dans l'item « TNM Regional Nodes Category Profile »
-------------------------------------	---

## 8.11 cotation du M

Nom de l'item	cotation du M
Numéro de l'item	8.11
définition	La cotation du M (qui représente la présence éventuelle de métastases) dans la classification TNM (lorsque M est différent de 0)
Nomenclature utilisée	la classification TNM (diffère selon la version) Notre recommandation est d'utiliser un fichier qui liste l'ensemble des codes de la classification de l'UICC TNM (version en fonction de la date du CR médical)
valeur attendue	chaîne de caractères Liste de codes de la classification TNM (présente et passées)
Règle	Même règle que pour le T et le N.
caractère obligatoire ou optionnel	Optionnel
comparaison avec d'autres standards	Dans mCODE la cotation du N est dans l'item « TNM Distant Metastases Category »

## 8.12 Version du TNM

Nom de l'item	Numéro de version du TNM
Numéro de l'item	8.12
définition	Numéro de version de la classification TNM
Nomenclature utilisée	UICC
valeur attendue	Entier (exemple, 6, 7, 8)
Règle	<p>En théorie, la version du TNM utilisée correspond à la version en vigueur lors du diagnostic de cancer.</p> <p>Par exemple, une tumeur dont le stade est réalisé en 2023 est codé dans la 8<sup>ème</sup> classification TNM</p> <p>La sixième édition était en application entre 2002 et 2009.</p> <p>La septième édition était en application entre en vigueur en 2010 jusqu'en 2018.</p>

	Si la version est précisée dans le CR (par exemple le CR d'anapath) on prend la version indiquée dans le compte-rendu.
caractère obligatoire ou optionnel	Optionnel

### 8.13 Type de TNM

Nom de l'item	Type de TNM
Numéro de l'item	8.13
définition	Le type de la classification TNM
Nomenclature utilisée	UICC
valeur attendue	c → clinique p → pathologique y →: après traitement néo-adjuvant
Règle	Le type est indiqué en fonction de la source de la donnée ou si c'est indiqué dans le compte rendu médical.
caractère obligatoire ou optionnel	Optionnel
Recommandations	Le TNM est codé au diagnostic. Il est invariant et unique (mono-valué).  Le TNM est clinique quand il est codé par le clinicien Le TNM est pathologique quand il est codé par le pathologiste quand le patient est opéré. On parle de yTNM après traitement médical et lorsque le patient est opéré.
Options	
comparaison avec d'autres standards	Dans mCODE <a href="#">Cancer Stage Group Profile</a> <ul style="list-style-type: none"> <li>- la stadification clinique est représentée par un préfixe « c », faisant référence à la stadification effectuée avant le premier traitement, ou à la stadification effectuée après le traitement néoadjuvant, mais avant la chirurgie VS 21905-5 (LOINC) VS 106248000 (SNOMED)</li> <li>- La classification pathologique, représentée par le préfixe « p », est utilisée pour les patients si la chirurgie est le premier traitement définitif, ou après un traitement néoadjuvant et une chirurgie post-néoadjuvante planifiée. VS 21899-0 (LOINC) VS 106249008 (SNOMED)</li> <li>- Le code « Autre » est censé inclure la stadification au moment de la récurrence ou de la progression (préfixe « r »), ou pour les cancers non reconnus auparavant et trouvés comme une découverte fortuite à l'autopsie (préfixe « a »). « other » (LOINC)</li> </ul>

## 8.14 Type de Stade

Nom de l'item	Type de Stade
Numéro de l'item	8.14
définition	Le système de classification ayant permis de définir le stade de la tumeur.
Nomenclature utilisée	Il sera nécessaire de fournir une proposition de liste complète et préciser qu'il serait préférable pour assurer l'interopérabilité que les utilisateurs ne peuvent ajouter de classifications et qu'il serait nécessaire d'assurer l'alignement vers les classifications choisies.  Il faudra identifier le type de stade par pathologie.
valeur attendue	Liste fermée (A compléter)
Règle	On prend le type (Ex : FIGO) quand il est renseigné dans le CR. Sinon on prend le type UICC TNM 8
caractère obligatoire ou optionnel	Optionnel

## 8.15 Stade de la tumeur (extension)

Nom de l'item	Stade de la tumeur (extension)
Numéro de l'item	8.15
définition	Valeur du stade suivant le système de classification utilisé.
Nomenclature utilisée	Liste fermée : 0 (cancers in situ), I, II, III, IV
valeur attendue	Chaîne de caractères
Règle	
caractère obligatoire ou optionnel	Optionnel
Recommandations	Si le stade est mentionné dans le CR médical, on récupère l'information structurée ou telle quelle sans recalcul à partir du TNM.

## 8.16 Type de grade histologique

Nom de l'item	Type de grade (histologique)
---------------	------------------------------

Numéro de l'item	8.16
définition	Le système de classification ayant permis de définir le grade de la tumeur.
Nomenclature utilisée	Code Snomed CT  Liste fermée listée ci-dessous
valeur attendue	Chaîne de caractères
caractère obligatoire ou optionnel	Optionnel
Recommandations	<p>Le grade du cancer (ou grade histologique) est une mesure de la différenciation des cellules et de l'agressivité de la tumeur.</p> <p>Par exemple, pour le cancer du sein, le pathologiste évalue trois paramètres: l'architecture tumorale, la forme et la taille du noyau de la cellule et l'activité mitotique. Le grade est exprimé par les lettres SBR, suivies des chiffres I, II ou III.</p> <p>Le système de classification du grade pour les cancers de la prostate est différent et est basé sur le score de Gleason.</p> <p>On attribue habituellement au grade un chiffre de 1 à 3 ou 4.</p> <p>Le grade peut aussi être exprimé en termes de « bas grade » pour les tumeurs les moins agressives et « haut grade » pour les tumeurs les plus agressives.</p>

Tumeur	classification	Code Snomed CT
Sein	Scarff Bloom Richarson (SBR) (aussi appelé Nottingham ou Elston-Ellis)	449205006
Prostate	Gleason, grade histopronostique ISUP	106241006
Sarcomes	FNCLCC	426757001
système nerveux central	OMS CNS	277460003
Rein	Fuhrman, <a href="#">échelle ISUP</a>	396192007
Tumeurs mentionnées dans les dossiers médicaux « haut grade », « bas grade », « grade intermédiaire »	OMS classification	1149429006
Si aucuns des 6 précédents	Grade Histologique	370114008

## 8.17 Grade histologique

Nom de l'item	Valeur du grade histologique
Numéro de l'item	8.17
définition	Valeur du grade histologique suivant le système de classification utilisé.
Nomenclature utilisée	Liste fermée à compléter
valeur attendue	Il sera nécessaire de fournir une proposition de liste complète et préciser qu'il serait préférable pour assurer l'interopérabilité que les utilisateurs ne peuvent ajouter de classifications et qu'il serait nécessaire d'assurer l'alignement vers les classifications choisies.
caractère obligatoire ou optionnel	Optionnel
Recommandations	Il est proposé, pour chaque type de tumeur (sein, prostate, etc.) une liste des systèmes de grade utilisé

## 9 - Evènements tumoraux

### Une récurrence locale est définie par :

- Une date de diagnostic (ou de découverte ou de premier traitement)
- Une anat-path (pour confirmer la récurrence)
- La même localisation tumorale
- Une latéralité (homolatérale à la tumeur primitive)
- Des biomarqueurs éventuels
- La même occurrence que la tumeur primitive

*Exemple : janvier 2015, récurrence locale d'un adénocarcinome du sein droit, RO -, RP -, HER2-*

### Les récurrences régionales sont définies par :

- Une date de diagnostic (ou de découverte ou de premier traitement)
- Une anat-path (pour confirmer la récurrence)
- Une localisation tumorale régionale par rapport à la tumeur primitive
- Un envahissement ganglionnaire homolatérale à la tumeur primitive
- Des biomarqueurs éventuels
- La même occurrence que la tumeur primitive

*Exemple : janvier 2020, récurrence ganglionnaire axillaire droite d'un adénocarcinome du sein droit, RO -, RP -, HER2 -*

### Les récurrences métastatiques sont définies par :

- Une date de diagnostic (ou de découverte ou de premier traitement)
- Une analyse anatomopathologique (pour confirmer la maladie métastatique) éventuelle
- Une ou plusieurs localisations métastatiques (souvent confirmée-s par imagerie)
- Des biomarqueurs éventuels
- La même occurrence que la tumeur primitive

*Exemple : janvier 2023, apparition de métastases osseuses et pulmonaires d'un adénocarcinome du sein droit traitées par chimiothérapie*

### 9.1 Type de l'évènement tumoral

Nom de l'item	Type de l'évènement tumoral
Numéro de l'item	9.1
définition	Les évènements tumoraux sont définis par les évènements cliniques suivants : pour les cancers solides :

	<ul style="list-style-type: none"> <li>- la récurrence locale,</li> <li>- la récurrence régionale</li> <li>- la métastase</li> </ul> <p>pour les cancers hématologiques :</p> <ul style="list-style-type: none"> <li>- la rechute</li> <li>- l'acutisation</li> </ul> <p>Nous définissons ces événements par leur première survenue. Cette première survenue correspond au début de l'évènement.</p>
Nomenclature utilisée	Chaîne de caractère
valeur attendue	Récurrence locale Récurrence régionale Métastase Rechute (cancers hémato) Acutisation (cancer hémato)
Règle	<p>Cette information étant rarement structurée, elle peut l'être à partir de différentes sources dont les comptes rendus médicaux.</p> <p>Il faut, pour une alimentation automatique, un algorithme qui définisse de façon précise :</p> <ul style="list-style-type: none"> <li>- une métastase par exemple quand c'est un cancer du poumon et qu'il y a une lésion sur un ganglion cervical</li> <li>- une récurrence régionale par exemple quand c'est un cancer du poumon à petites cellules avec un ganglion sus-claviculaire homo-latéral</li> <li>- une récurrence locale pour un cancer ORL avec l'apparition d'un ganglion cervical</li> <li>- Etc...</li> </ul> <p>Et cela pour l'ensemble des cancers.</p> <p>Variable mono-évaluée</p>
caractère obligatoire ou facultatif	Optionnel
comparaison avec d'autres standards	Voir la comparaison avec OMOP en annexe

Pour chaque événement tumoral, (sauf pour la tumeur primaire), on propose de distinguer 3 dates :

- La date de première évocation de l'évènement dans le dossier (clinique ou imagerie)
- La date de diagnostic histologique
- La première date de traitement de l'évènement (cf. Chapitre traitement)

## 9.2 Date de première évocation de l'évènement tumoral

Nom de l'item	Date de première évocation de l'évènement tumoral
Numéro de l'item	9.2
Définition	Date de première évocation de l'évènement dans le dossier médical.
Nomenclature utilisée	
valeur attendue	YYYY - MM - DD
Règle	<p>Pour la gestion des dates de début des évènements tumoraux, nous avons déjà vu plus haut que ces évènements sont définis par leur première occurrence.</p> <p>Celle-ci marque le commencement ou la date de l'évènement tumorale</p> <ul style="list-style-type: none"> <li>- récurrence locale (1<sup>er</sup> occurrence)</li> <li>- récurrence régionale (1<sup>er</sup> occurrence)</li> <li>- maladie métastatique (1<sup>er</sup> occurrence)</li> </ul> <p>Cf annexe - évolution du cancer</p> <p>Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.</p>
justification du choix de la règle (exemple)	<p>Un consensus sur une seule date est difficile car un cancer, ou tout évènement tumoral (comme toute maladie chronique) ne débute jamais à une date donnée contrairement à une maladie aiguë comme une pneumonie à pneumocoques ou un infarctus du myocarde.</p> <p>En fonction des questions posées, il sera possible d'utiliser l'une ou l'autre date.</p>
caractère obligatoire ou facultatif	Optionnel (cf schéma (2))
Recommandations	
Options	
comparaison avec d'autres standards	<p>mCODE</p> <p>La date à laquelle l'existence de la condition a été affirmée ou reconnue pour la première fois.</p> <p>OMOP (voir annexe)</p>



### 9.3 Date de diagnostic histologique

Nom de l'item	Date de diagnostic histologique
Numéro de l'item	9.3
Définition	Date de confirmation du diagnostic histologique.
Nomenclature utilisée	
valeur attendue	YYYY - MM - DD
Règle	Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
justification du choix de la règle (exemple)	
caractère obligatoire ou facultatif	Optionnel
Recommandations	
Options	
comparaison avec d'autres standards	

### 9.4 Date de début du traitement de l'évènement tumoral

Nom de l'item	Date de début du traitement de l'évènement tumoral
Numéro de l'item	9.4
définition	Date du premier traitement après détection de l'évènement (date de diagnostic histo ou date de première évocation).
valeur attendue	YYYY - MM - DD
Règle	Premier traitement après détection de l'évènement (date de diagnostic histo ou date de première évocation).  Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
caractère obligatoire ou facultatif	Optionnel
Recommandations	Quand le traitement est réalisé, la date de début de traitement est la date la plus fiable car il s'agit le plus souvent d'une information

	structurée - date du CR de chirurgie - date de 1 <sup>er</sup> jour de chimiothérapie - date de début de radiothérapie
--	---

## 9.5 Code histologique/ morphologique

Nom de l'item	Code histologique/ morphologique
Numéro de l'item	9.5
définition	Description morphologique de l'évènement tumoral
Nomenclature utilisée	CIM-O-3 Morpho
valeur attendue	chaîne de caractères (ex:8144/6)
Règle	Notre recommandation pour ce qui est du code morphologique, est de le prendre dans le compte-rendu d'anapath lorsqu'il y a une biopsie.
caractère obligatoire ou facultatif	Optionnel
Recommandations	<p>Quand la maladie progresse, de plus en plus souvent, l'histologie confirme ou permet de déterminer le traitement.</p> <p>Point d'attention, sur le fait qu'une même tumeur peut avoir plusieurs codes morphologiques, d'autant que les tumeurs peuvent être hétérogènes (exemple : association de carcinome infiltrant et de carcinome in situ).</p> <p>Au cours de l'évolution d'une même tumeur, la morphologie est parfois modifiée même si on conclut qu'il s'agit d'une évolution de la tumeur initiale.</p> <p>Néanmoins, d'un point de vue clinique on s'appuie sur la morphologie pour faire un diagnostic de l'évolution d'un cancer d'une tumeur initiale par rapport à l'apparition d'un évènement tumoral comme par exemple un second cancer (exemple : si apparait un adénocarcinome au niveau du foie après le diagnostic d'un adénocarcinome du poumon, c'est que probablement nous sommes face à une métastase du cancer du poumon et non face à un deuxième cancer.</p>
Options	Pour une même tumeur, il est souvent retrouvé plusieurs codes anapath (histo/morpho) ; nous aurions pu faire le choix de collecter une donnée multivariée. Cette option n'a pas été retenue.
comparaison avec d'autres standards	Pour les cancers secondaires, le suffixe comportemental de la CIM-O-3 doit être /6.

## 9.6 Code topographique des métastases

Nom de l'item	Code topographique des métastases
Numéro de l'item	9.6
définition	Localisation des métastases dans l'évènement « maladie métastatique ».
Nomenclature utilisée	CIM-O-3 Topo
valeur attendue	chaîne de caractères (ex:C18.4)
Règle	Cette information ne concerne que les métastases de l'évènement « maladie métastatique » ; on recueille le code topographique des métastases ainsi que leur date d'apparition (ou de découverte).
caractère obligatoire ou facultatif	Optionnel mais obligatoire dans le cas d'évènements métastatiques
Recommandations	Les localisations métastatiques ne nécessitent pas la même précision que la localisation des tumeurs primitives. Donc on peut se limiter au premier niveau de topographie de la classification CIM-O-3 ex : C22 Foie C34 Poumon C71 cerveau

## 9.7 Dates de découverte des métastases par topographie

Nom de l'item	Date de découverte des métastases par topographie
Numéro de l'item	9.7
définition	Date de survenue des métastases rattachées à un code topographique de la métastase et à l'évènement « maladie métastatique ».
valeur attendue	YYYY - MM - DD
Règle	Cette information ne concerne que les métastases de l'évènement « maladie métastatique ». Dans cet évènement, il faut recueillir les codes topographiques des métastases ainsi que leur date d'apparition au fur et à mesure de leur survenue.  Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
caractère obligatoire ou facultatif	Optionnel

## 9.8 Investigations réalisées pour établir le diagnostic des évolutions tumorales

Nom de l'item	Investigation réalisées
Numéro de l'item	9.8
définition	Méthode d'investigation et de confirmation du diagnostic de l'événement tumoral
Règle	Dès qu'une investigation avec un ordre plus élevé (cf. infra) est réalisée, cette modalité prime sur les codes inférieurs.
valeur attendue	1 clinique 2 investigation clinique (y compris imagerie) 3 biomarqueurs spécifiques 4 pour cytologie 5 pour histologie 6 pour les tests cytogénétiques 0 pour inconnu
caractère obligatoire ou facultatif	Optionnel
Recommandations	Référentiel : <a href="https://publications.iarc.fr/Book-And-Report-Series/Iarc-Technical-Publications/Standards-And-Guidelines-For-Cancer-Registration-In-Europe-2003">https://publications.iarc.fr/Book-And-Report-Series/Iarc-Technical-Publications/Standards-And-Guidelines-For-Cancer-Registration-In-Europe-2003</a>

## Le standard OMOP

Trois types d'épisodes sont introduits dans OMOP :

1. Épisode global de la maladie : couvre toute la durée de la maladie. L'épisode global de la maladie est l'épisode parent pour les autres épisodes de la maladie, l'étendue de la maladie et la dynamique de la maladie et est toujours présent. Dans notre modèle, cela correspondant au cancer.
2. Étendue de la maladie : décrit jusqu'où le cancer s'est propagé. L'épisode commence le jour du modificateur diagnostique indiquant une nouvelle étendue de la maladie et se termine au début d'un épisode de traitement ou au jour d'un modificateur diagnostique indiquant une étendue différente. L'étendue de la maladie comprend :
  - Maladie confinée : Maladie confinée à un organe ou à un site (pT1-pT4) sans métastases ganglionnaires (régionales) ou métastatique.
  - Maladie régionale (invasive) : Maladie locale/régionale-confinée avec métastases ganglionnaires (régionales) (tout pT et N1 ou plus). Commence avec le premier modificateur diagnostique positif de la classe conceptuelle « Invasion/extension » au-delà des organes ou des ganglions lymphatiques

régionaux et se termine le jour du premier modificateur positif de la classe conceptuelle « Métastases ».

- Maladie à distance (métastatique) : commence par le premier modificateur positif de la classe conceptuelle « Metastasis » et se termine au moment du décès ou de la dernière observation.
- 3. Dynamique de la maladie : décrit l'état du cancer (maladie stable, rémission, progression). L'épisode commence le jour du modificateur diagnostique indiquant une nouvelle dynamique et se termine le jour d'un modificateur diagnostique indiquant une dynamique différente ou un jour de traitement. La chirurgie commence un nouvel épisode de « rémission », sauf indication contraire dans le rapport de chirurgie.
  - Maladie stable
  - Rémission - On suppose que la chirurgie déclenche une période de rémission. Si l'on ne sait pas s'il y a une diminution ou une disparition des signes et symptômes du cancer, l'état « Rémission » peut être utilisé à la place de « Rémission complète » et « Rémission partielle ».
  - Rémission partielle - Dans la « rémission partielle ».
  - Rémission complète - Dans la « rémission complète », tous les signes et symptômes du cancer ont disparu.
  - Progression

Pour ce qui est des périodes de rémission et progression il y a souvent une alternance rémission/Progression. En revanche, une période de rémission commence toujours par une preuve clinique ou une chirurgie et se poursuit jusqu'à une autre preuve clinique ou un changement de traitement. En ce qui concerne l'étendue de la maladie, elle est considérée comme confinée jusqu'à une évolution loco-régionale ou métastatique.

## 10 - Etat général et examen physique (taille, poids, IMC, PS)

Ces différents items sont des mesures survenant à différents moments de l'histoire de la maladie cancéreuse.

Pour chaque item, ce sont des données numériques ou alpha-numériques qui se caractérisent par :

- Un type de mesure : poids, taille, indice de masse corporelle, Performance status
- Une valeur ; poids en kilogrammes, taille en centimètre, IMC, Performance status (code selon la classification OMS)
- Un attribut : Unité/Echelle de mesure
- Une date de la mesure

Ces quatre items sont tous rattachées au patient, avec un nom de variable (poids, taille, IMC, PS), une valeur et une date.

### 10.1 Type de mesures de l'état général et de l'examen physique

Nom de l'item	Nom de la variable mesurant l'état physique et l'état général
Numéro de l'item	10.1
Définition	Mesures sur l'état général du patient (PS) et son examen physique (taille, poids)
Valeur attendue de la mesure	PS (Performance Status) ; Poids ; Taille ; IMC brut
caractère obligatoire ou optionnel	optionnel

### 10.2 Valeurs de la mesure

Nom de l'item	Mesures sur l'état général
Numéro de l'item	10.2
Définition	Valeur de la mesure sur l'état général du patient (PS) et son examen physique (taille, poids)
Valeur attendue de la mesure	Chaines de caractères  PS : codes 0,1,2,3,4 Poids : entier en kilogramme Taille : entier en centimètres IMC : nombre flottant (ex :18,5)
Règle	La recommandation est de conserver toutes les mesures avec les dates. Cette recommandation concerne le PS, le poids, la taille et l'IMC.

	<p>Concernant le PS, et si les médecins codent en index de Karnofsky (IK), la recommandation est de le transformer en PS selon la classification OMS et d'utiliser la table de correspondance entre PS et IK :</p> <table border="1"> <thead> <tr> <th>Score ECOG</th><th>indice de Karnofsky</th></tr> </thead> <tbody> <tr> <td>0</td><td>100</td></tr> <tr> <td>1</td><td>80-90</td></tr> <tr> <td>2</td><td>60-70</td></tr> <tr> <td>3</td><td>40-50</td></tr> <tr> <td>4</td><td>10-30</td></tr> <tr> <td>5</td><td>0</td></tr> </tbody> </table> <p>Lorsque la donnée est disponible sous forme structurée dans une source (logiciel de prescription, dossier de soins), on priorise cette source par rapport à une source de données non structurée.</p> <p>La donnée peut être récupérée par un traitement du langage dans les documents médicaux ou dans les outils de prescription.</p>	Score ECOG	indice de Karnofsky	0	100	1	80-90	2	60-70	3	40-50	4	10-30	5	0
Score ECOG	indice de Karnofsky														
0	100														
1	80-90														
2	60-70														
3	40-50														
4	10-30														
5	0														
justification du choix de la règle (exemple)	Si un médecin code un PS avec une valeur ambiguë, (exemple 1-2), on conserve la valeur la plus « péjorative » c'est-à-dire le chiffre le plus grand.														
caractère obligatoire ou optionnel	Optionnel														

### 10.3 Date de recueil de la mesure

Nom de l'item	Date de recueil de la mesure
Numéro de l'item	10.3
Définition	La date à laquelle la mesure est recueillie
Valeur attendue de la mesure	YYYY - MM - DD
Règle	<p>On prend chaque mesure associée à une date.</p> <p>Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.</p>
caractère obligatoire ou optionnel	Optionnel

# LES TRAITEMENTS

Les traitements sont classés en quatre types :

- Les traitements chirurgicaux
- Les traitements médicaux ou systémiques à visée anti-cancéreuse (chimiothérapie, immunothérapie, thérapies ciblées, hormonothérapies ou autres traitements médicaux)
- La radiothérapie (y compris la curiethérapie, la protonthérapie et les autres traitements utilisant les radiations ionisantes)
- Les autres traitements en dehors des trois types précédents (radiologie interventionnelle, laser, cryothérapie, thérapies cellulaires, etc.).

Les traitements chirurgicaux, les traitements médicaux et les autres traitements sont décrits avec un nombre limité d'items.

Par contre, la radiothérapie est décrite de façon très précise dans le modèle dit MED-OSIRIS. Un lien est réalisé entre la radiothérapie décrite dans le modèle OSIRIS RWD clinique et le modèle OSIRIS radiothérapie (OSIRIS\_CT Radiothérapie).

Des groupes de réflexion se constituent pour décrire de façon plus précise la chirurgie et les traitements de chimiothérapie.

Ces travaux seront intégrés ultérieurement.

Un traitement est défini par :

- Un type de traitement (chirurgie, traitement médical, radiothérapie, autre)
- Une date de traitement (ou une date de début de traitement)
- Une précision concernant le type de traitement (type d'acte chirurgical, molécules utilisées, technique de radiothérapie, etc.)
- Un lieu de réalisation de l'acte en différenciant un acte réalisé dans l'établissement d'un acte effectué en dehors, dans une autre structure hospitalière. En général, lorsque le traitement est réalisé dans l'établissement, l'information est structurée à la source et est présente dans la source PMSI (hospitalisation, actes et séances) ; lorsque le traitement est réalisé en dehors, l'information peut être structurée par NLP à partir d'un compte rendu médical.

L'information sur la réalisation de l'acte peut être intégrée dans le modèle pour tous les items où cela sera considéré comme nécessaire.

L'information concernant la visée du traitement avec des modalités comme « traitement curatif » ou « palliatif » n'a pas été retenue. Par contre, l'information sur les essais cliniques et les accès précoces est conservée.

L'information concernant la date de fin de traitement n'est pas retenue dans le minimum data set ; elle pourra être recalculée si nécessaire.



# 11 - Chirurgie

## 11.1 Date de la chirurgie

Nom de l'item	Date de la chirurgie
Numéro de l'item	11.1
Définition	Date de l'acte chirurgical
valeur attendue	YYYY - MM - DD
Règle	<p>La date de la chirurgie est la date de la réalisation de l'acte.</p> <p>Si la date ne peut être récupérée, c'est la date du compte-rendu opératoire.</p> <ul style="list-style-type: none"><li>- YYYY - MM - DD si la source est le compte rendu opératoire (CRO)</li><li>- YYYY - MM - DD si l'information est récupérée à partir des données du PMSI dans un résumé de sortie (car la date précise de la chirurgie est précisée dans le résumé d'unité médicale)</li></ul> <p>Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.</p>
justification du choix de la règle (exemple)	<p>Si l'information est récupérée à partir du PMSI, on prend la date de l'acte qui figure dans le résumé d'unité médicale (RUM) ; c'est le cas lorsque l'acte a été fait dans l'établissement lui-même.</p> <p>Si la source n'est pas le PMSI comme c'est le cas où l'acte chirurgical a été fait dans un autre établissement, la date de l'acte chirurgical doit être récupérée dans le CRO récupéré de l'établissement extérieur à l'établissement.</p>
caractère obligatoire ou optionnel	Optionnel
Recommandations	<p>La recommandation est de récupérer toutes les dates de tous les actes chirurgicaux du patient.</p> <p>Plusieurs sources de données sont possibles pour obtenir l'information. On propose l'ordre suivant :</p> <ul style="list-style-type: none"><li>- Les données du PMSI (information structurée)</li><li>- Le compte-rendu opératoire (information textuelle à structurer)</li><li>- Les autres comptes rendus médicaux</li></ul>

## 11.2 Type de chirurgie

Nom de l'item	Type de chirurgie
Numéro de l'item	11.2

définition	Description de l'acte chirurgical
Nomenclature utilisée	Code CCAM
Valeur attendue	Liste fermée à fournir
Règle	<p>A partir d'une liste restreinte sur les codes CCAM en filtrant sur les actes chirurgicaux liés aux traitements.</p> <p>On conserve les actes chirurgicaux non thérapeutiques (chirurgies reconstructrices ou prophylactiques)</p>
justification du choix de la règle (exemple)	<p>La classification commune des actes médicaux (CCAM) est une classification française utilisée pour la facturation et la description de l'activité par les hôpitaux. Elle est mise à jour régulièrement, parfois avec retard sur l'état de l'art et sur les innovations.</p> <p>La CCAM est la liste des actes médicaux techniques, utilisant des codes, commune aux secteurs privé et public, qui permet la description de l'activité chirurgicale.</p> <p>La CCAM a remplacé progressivement le CdAM pour le PMSI dans les établissements publics et privés en 2004.</p> <p>Il existe des actes dits non chirurgicaux ou d'imagerie mais il est possible de filtrer les actes chirurgicaux des autres actes.</p>
caractère obligatoire ou optionnel	Optionnel
Recommandations	<p>Concernant l'imagerie interventionnelle, la recommandation est de classer l'imagerie à visée interventionnelle qui est codée également dans les codes CCAM dans « les autres traitements ».</p> <p>Les 4 premières lettres de chaque code sont significatives</p> <p>Concernant la source de l'item, la recommandation est de bien distinguer les sources d'information concernant cet item et de les stocker dans un item « Acte chirurgical réalisé dans l'établissement ou en dehors ».</p> <p>La préconisation est de créer un item « source de l'information type de chirurgie ».</p> <p>Quel que soit la source PMSI (données structurées) ou les CR extérieurs océrés via un algo de NLP, le code CCAM est récupérable.</p> <p>Un acte chirurgical est donc souvent multivarié avec plusieurs codes CCAM codés dans la même chirurgie. La recommandation est d'agréger les codes CCAM utilisés pour décrire la chirurgie sur une même date.</p> <p>Un acte est souvent codé par plusieurs codes. Ceci nécessite de choisir une liste de codes parmi les codes possibles.</p>

	Un même code est en double car il y a à la fois l'activité 1 (chirurgien) et l'activité 4 (anesthésiste) qui est codée, ce qui provoque l'existence de doublons. Un seul code sera retenu par couple (date,patient).
Options	On n'agrège pas les codes et on aura une ligne par code CCAM avec une date. Plusieurs actes CCAM peuvent figurer pour une même date de chirurgie du CRO. Cela consiste à créer plusieurs lignes avec les mêmes dates de chirurgie.
comparaison avec d'autres standards	OMOP Snomed CT, CPT Classification ICD-10-PCS

### 11.3 Chirurgie réalisée dans l'établissement

Nom de l'item	Chirurgie réalisée dans l'établissement
Numéro de l'item	11.3
Définition	Indicateur du lieu de réalisation de la chirurgie
Nomenclature utilisée	Booléen
valeur attendue	OUI / NON
caractère obligatoire ou optionnel	Obligatoire
Remarques	Lorsque l'acte est réalisé dans l'établissement, la source de données pour la chirurgie sera le PMSI de l'établissement. Sinon, la chirurgie est repérée par NLP dans des CR.
comparaison avec d'autres standards	OMOP précise le prestataire de soin qui a réalisé la chirurgie mCODE a un item sur la localisation de la chirurgie

## 12 - Traitements médicaux

### 12.1 Ligne de traitement médical

Nom de l'item	Ligne de traitement
Numéro de l'item	12.1
définition	Par convention, on réserve la notion de ligne de traitement aux patients en phase métastatique. Ainsi, on compte le nombre de lignes à partir de la première administration en phase métastatique.
valeur attendue	Entier (Numéro de ligne de traitement)

Règle	Les règles, options et recommandations figurent dans un texte en annexe
caractère obligatoire ou optionnel	Optionnel

## 12.2 Code de la molécule

Nom de l'item	Code de la molécule
Numéro de l'item	12.2
définition	<p>Le code de la molécule correspond à un code ATC qui renvoie à la Dénomination commune internationale (DCI).</p> <p>La DCI d'un médicament correspond au nom scientifique de la molécule (ou principe actif) responsable de l'effet thérapeutique. Créée par l'Organisation mondiale de la santé, elle est commune aux pays du monde entier.</p> <p>L'information devrait être présente sous cette forme dans les systèmes d'information hospitaliers car depuis le 1er janvier 2015, les médecins sont tenus de prescrire toute spécialité pharmaceutique en DCI.</p> <p>Dans le cas contraire, il sera nécessaire de transformer le nom de Spécialité d'un médicament en DCI</p>
Nomenclature utilisée	Code ATC
valeur attendue	<p>Exemples :</p> <ul style="list-style-type: none"> <li>- L01XC18 (pembrolizumab)</li> <li>- L01CD01 (paclitaxel)</li> <li>- L01XY02 : association fixe de pertuzumab et trastuzumab</li> </ul>
caractère obligatoire ou optionnel	Optionnel

## 12.3 Nom de la molécule administrée hors ATC

Nom de l'item	Nom de la molécule administrée hors ATC
Numéro de l'item	12.3
Définition	Nom d'une molécule hors ATC. C'est souvent le cas des molécules en phase d'essais cliniques non commercialisée.
Nomenclature utilisée	Dénomination commune internationale (DCI).

valeur attendue	Chaine de caractère
Règle	Quand la molécule administrée ne fait pas partie de la classification ATC et qu'elle est enregistrée dans le logiciel de prescription type CHIMIO, alors on la prend telle quelle.
justification du choix de la règle (exemple)	
caractère obligatoire ou optionnel	Optionnel
Recommandations	Les médicaments en accès précoces peuvent être inclus dans cette variable.

## 12.4 Traitement médical administré dans l'établissement

Nom de l'item	Traitement administré dans l'établissement
Numéro de l'item	12.4
définition	Indicateur du lieu d'administration du traitement médical
valeur attendue	OUI / NON
caractère obligatoire ou optionnel	Obligatoire

## 12.5 Date de début du traitement médical

Nom de l'item	Date
Numéro de l'item	12.5
Définition	Date de début du traitement de la molécule administrée
valeur attendue	YYYY - MM - DD  Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
caractère obligatoire ou optionnel	Optionnel
Recommandations	Date d'administration pour les chimiothérapies et immunothérapies  Date de prescription pour les traitements oraux (thérapies ciblées, hormonothérapie)

## 12.6 Essai clinique

Nom de l'item	Essai clinique
Numéro de l'item	12.6
Définition	Le traitement s'inscrit dans le contexte d'un essai clinique
valeur attendue	Oui/Non
caractère obligatoire ou optionnel	Optionnel

## 12.7 Médicament en accès précoce

Nom de l'item	Accès précoce
Numéro de l'item	12.7
Définition	Le traitement s'inscrit dans le contexte d'un accès précoce
valeur attendue	Oui/Non
caractère obligatoire ou optionnel	Optionnel

# 13 - Radiothérapie

Note : les variables présentes dans ce chapitre seront amenées à évoluer pour correspondre aux variables minimales permettant d'identifier une observation dans le modèle MED-OSIRIS RADIOTHERAPIE.

## 13.1 Type de traitement de radiothérapie

Nom de l'item	Type de traitement de radiothérapie
Numéro de l'item	13.1
Définition	Un traitement de radiothérapie peut être administré en une seule séance (radio-chirurgie) soit une suite de séance de radiothérapie.
Nomenclature utilisée	Radiothérapie Curiethérapie Protonthérapie Radiothérapie stéréotaxique Irradiation corporelle totale Autre (Nomenclature utilisée en radiothérapie au format DICOM RT)

caractère obligatoire ou optionnel	Optionnel
Recommandations	<p>Le système doit gérer les interruptions de traitement qui peuvent survenir de façon protocolaire (association radio-chimiothérapie) soit pour un arrêt en lien avec la survenue d'effets indésirables.</p> <p>On rappelle que RT-Plan contient les infos curie, photon, électrons, protons. Les informations sur la stéréotaxie, le per-op ou l'ICT peuvent être déduites de la dose</p>

## 13.2 Traitement de radiothérapie réalisé dans l'établissement

Nom de l'item	Traitement de radiothérapie réalisé dans l'établissement
Numéro de l'item	13.2
Définition	Indicateur du lieu de réalisation de la radiothérapie
Nomenclature utilisée	Booléen
valeur attendue	OUI/NON
caractère obligatoire ou optionnel	Optionnel
Recommandations	<p>Le système peut se baser sur l'information du PMSI c'est-à-dire sur la source de données. Si le traitement est réalisé dans l'établissement, il existe un RSS dit « séance ».</p> <p>En l'absence de données issues du PMSI, la radiothérapie peut être considérée en dehors de l'établissement.</p> <p>Un traitement de radiothérapie débuté est toujours poursuivi dans le même établissement.</p>

## 13.3 Topographie de l'organe irradié

Nom de l'item	Topographie de l'organe irradié
Numéro de l'item	17.3
Définition	Codes décrivant les topographies du corps où les traitements de radiothérapie ont été délivrés.
Nomenclature utilisée	CIM-10
caractère obligatoire ou optionnel	Optionnel

Recommandations	L'information est à récupérer en priorité dans les RSS « séances » du PMSI et en l'absence d'information structurée, dans les compte rendus médicaux
Options	Un code pour la radiothérapie du corps entier est nécessaire

### 13.4 Date de début de la radiothérapie

Nom de l'item	Date de début de la radiothérapie
Numéro de l'item	17.4
Définition	Date de début du traitement de radiothérapie
valeur attendue	YYYY - MM - DD  Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
Recommandations	Les dates de traitement se retrouvent dans RT-RECORD
caractère obligatoire ou optionnel	Optionnel

### 13.5 Lien vers le set de données radiothérapie V1.0

Nom de l'item	Lien vers le set de données radiothérapie V1.0
Numéro de l'item	17.5
Définition	Lien avec le DICOM RT du patient  Dans le cas où le traitement est réalisé dans l'établissement, il est recommandé d'alimenter la source OSIRIS RADIOTHERAPIE pour récupérer de façon automatique les données au format DICOM RT
Règle	On récupère dans le dossier DICOM RT du patient ce qui a été délivré avec la date de début du traitement.
caractère obligatoire ou optionnel	Optionnel



## 14 - Autres traitements du cancer

### 14.1 Type de traitement

Nom de l'item	Type de traitement réalisé
Numéro de l'item	14.1
Définition	Un autre traitement réalisé est un traitement qui n'est pas chirurgical, médical (chimiothérapie par exemple) ou une radiothérapie
Nomenclature utilisée	LOINC
valeur attendue	Liste fermée Ex : cryothérapie, laser, cimentoplastie
caractère obligatoire ou optionnel	Optionnel
Recommandations	LOINC doit être utilisé mais il sera possible de restreindre la liste de valeurs à des valeurs couramment utilisées.

### 14.2 Date de réalisation de l'acte (autre traitement)

Nom de l'item	Date de réalisation de l'acte (autre traitement)
Numéro de l'item	14.2
Définition	Date de réalisation de l'acte (autre traitement)
valeur attendue	YYYY - MM - DD Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
caractère obligatoire ou optionnel	Optionnel

### 14.3 Lieu de réalisation de l'acte (autre traitement)

Nom de l'item	Lieu de réalisation de l'acte (autre traitement)
Numéro de l'item	14.3
définition	L'acte (autre traitement) a été réalisé dans l'établissement
Nomenclature utilisée	Oui/Non
caractère obligatoire ou optionnel	Optionnel

## 15 - IMAGERIE ET MÉDECINE NUCLÉAIRE

Les variables présentes dans ce chapitre seront amenées à évoluer pour correspondre aux variables minimales permettant d'identifier une observation dans le modèle MED-OSIRIS.

### 15.1 Date de l'examen d'imagerie

Nom de l'item	Date de réalisation ou d'acquisition de l'examen d'imagerie
Numéro de l'item	15.1
Définition	Date et heure de réalisation ou d'acquisition de l'examen d'imagerie
valeur attendue	YYYY - MM - DD
Règle	<p>La date de l'acte d'imagerie est la date de la réalisation de l'acte.</p> <p>Si la date ne peut être récupérée à partir d'une source structurée (PACS), c'est la date du compte-rendu d'imagerie qui prévaut</p> <ul style="list-style-type: none"><li>- YYYY - MM – DD et HH:MM car nécessaire pour identifier une suite d'images si l'information est récupérée à partir des données du PACS</li><li>- YYYY - MM - DD si la source est le compte rendu</li></ul> <p>Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.</p>
caractère obligatoire ou facultatif	Obligatoire
Recommandations	<p>Si la source n'est pas le PACS, si c'est le cas où l'examen d'imagerie a été fait dans un autre établissement ou un cabinet privé, la date de l'examen est récupérée dans le compte-rendu récupéré de l'établissement extérieur à l'établissement.</p> <p>La recommandation est de récupérer toutes les dates de tous les examens d'imagerie du patient.</p> <p>Plusieurs sources de données sont possibles pour obtenir l'information. On propose l'ordre suivant :</p> <ul style="list-style-type: none"><li>- Les données du PACS (information structurée)</li><li>- Les compte-rendus d'examens d'imagerie (information textuelle à structurer)</li><li>- Les autres comptes-rendus médicaux</li></ul> <p>La recommandation est d'extraire la date des données DICOM de l'image, ceci étant plus fiable que les comptes rendus médicaux.</p>

## 15.2 Modalité utilisée pour l'examen

Nom de l'item	Modalité à l'origine de l'examen d'imagerie
Numéro de l'item	15.2
Définition	Type d'équipement utilisé permettant l'acquisition des images
Nomenclature utilisée	Code CCAM
valeur attendue	Liste de l'ensemble des modalités utilisées (A préciser)
justification du choix de la règle (exemple)	<p>Double avantage d'utiliser le code CCAM :</p> <ul style="list-style-type: none"> <li>- Contient l'information sur la topographie.</li> <li>- La métadonnée recueillie dans les fichiers DICOM ne respecte pas une classification standardisée en raison du caractère permissif des normes DICOM</li> </ul>
caractère obligatoire ou facultatif	Obligatoire
Recommandations	<p>La modalité issue du code CCAM fonctionne en France mais n'est pas aligné sur les standards internationaux.</p> <p>Il existe un risque d'incohérence entre les données DICOM et le code CCAM.</p> <p>La recommandation est l'utilisation du DICOM pour automatiser le remplissage de cet item en récupérant l'information depuis le tag DICOM correspondant.</p>
Options	Une option est d'utiliser le code CCAM de l'examen qui contient à la fois l'information sur la modalité et la topographie de l'étude.

## 15.3 Lien vers le set de données du modèle OSIRIS imagerie

Nom de l'item	Lien vers le set de données imagerie
Numéro de l'item	18.3
Définition	Numéro unique permettant d'identifier l'étude
valeur attendue	Chaine de Caractères Dicom Tag (0020,000D)
Règle	Nous récupérons le numéro unique du PACS lorsque l'image est réalisée dans l'établissement ou lorsqu'elle est acquise à l'extérieur et intégrée au PACS. Dans ce dernier cas, les informations disponibles sont identiques à celles d'une image capturée dans l'établissement.
justification du choix de la règle (exemple)	La différence entre l'étude et la série réside dans le fait que l'étude, par exemple un scanner thoraco-abdomino-pelvien, englobe l'ensemble de l'examen, tandis que les séries désignent les différentes parties de cet

	<p>examen.</p> <p>Par exemple, une série peut être le scanner thoracique, une autre série peut être le scanner abdomino-pelvien sans injection, et une troisième série peut être le scanner abdomino-pelvien avec injection dans un même examen (étude).</p> <p>Ainsi, toutes les séries regroupées dans l'étude sont nécessaires pour obtenir une vision complète de l'examen.</p> <p>Ensuite, les images spécifiques sont répertoriées dans chaque série.</p>
caractère obligatoire ou facultatif	Optionnel

## 16 - INFORMATION CONCERNANT LES ÉCHANTILLONS BIOLOGIQUES

Informations permettant de décrire les échantillons utilisés et conservés pour être ré-utilisés dans un projet de recherche.

### 16.1 Lieu de conservation de l'échantillon

Nom de l'item	Lieu de conservation de l'échantillon
Numéro de l'item	16.1
Définition	<p>Il est important de faire une distinction entre les échantillons conservés dans le laboratoire d'anatomopathologie de l'établissement et ceux stockés au Centre de Ressources Biologiques (CRB).</p> <p>Ces deux entités utilisent des logiciels différents et constituent ainsi deux sources d'informations distinctes.</p> <p>Généralement, les échantillons de type bloc histologique et lame sont conservés dans le laboratoire d'anatomopathologie, tandis que les échantillons congelés et autres types sont entreposés au CRB.</p>
Nomenclature utilisée	Liste fermée
valeur attendue	<ul style="list-style-type: none"> <li>- Laboratoire d'anatomopathologie de l'établissement</li> <li>- CRB de l'établissement</li> <li>- Hors de l'établissement</li> </ul>
Règle	Cette métadonnée permet d'identifier la source de la donnée structurée

	du laboratoire d'anatomopathologie de l'établissement et/ou du CRB du de l'établissement.
caractère obligatoire ou optionnel	Optionnel

## 16.2 Identifiant de l'échantillon biologique

Nom de l'item	Identifiant de l'échantillon biologique
Numéro de l'item	16.2
Définition	Identifiant de l'échantillon dans l'établissement (lien avec le set de données échantillon biologique)
valeur attendue	Chaîne de caractères
Règle	NA
justification du choix de la règle (exemple)	NA
caractère obligatoire ou optionnel	Obligatoire
Recommandations	<p>Seuls les échantillons pris en charge dans l'établissement où le patient est suivi sont structurés dans la base OSIRIS RWD, ce qui inclut les échantillons du laboratoire d'anatomopathologie de l'établissement et du Centre de Ressources Biologiques (CRB) affilié a l'établissement de prise en charge.</p> <p>Cette approche présente l'avantage de maintenir une information à jour sur la disponibilité des échantillons</p>
Options	Les échantillons qui viendraient d'un CRB extérieur a l'établissement sont rarement connus et sont souvent des documents statiques. Cette option n'a pour cette raison pas été retenue.

## 16.3 Date de prélèvement de l'échantillon

Nom de l'item	Date de prélèvement de l'échantillon
Numéro de l'item	16.3
Définition	Date de prélèvement de l'échantillon
valeur attendue	YYYY - MM - DD
Règle	Cette donnée est collectée au moyen de la donnée structurée du laboratoire d'anapath et/ou du CRB de l'établissement.

	Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.
justification du choix de la règle (exemple)	NA
caractère obligatoire ou facultatif	Obligatoire

## 16.4 Nature de l'échantillon pour un CRB

Nom de l'item	Nature de l'échantillon
Numéro de l'item	16.4
Définition	Information indiquant la nature de l'échantillon.
Nomenclature utilisée	Unified Medical Language System UMLS)
valeur attendue	<p>Chaine de caractères</p> <p>Liste fermée : (A compléter)</p> <p>UMLS:C0005767 pour du sang</p> <p>UMLS:C0085983 pour une tumeur (bloc)</p> <p>UMLS:C1548793 pour un tissu congelé</p> <p>UMLS:C1519524 pour un tissu FFPE</p> <p>UMLS:C0042036 pour de l'urine</p> <p>UMLS:C0032105 pour du plasma</p> <p>UMLS:C0035668 pour du RNA</p> <p>UMLS:C0012854 pour de l'ADN</p> <p>UMLS:C1321301 pour une cellule du sang périphérique</p> <p>UMLS:C0439673 pour non connu</p> <p>UMLS :XXXXXXX pour des selles</p> <p>UMLS :XXXXXXX pour une poudre</p>
caractère obligatoire ou optionnel	Obligatoire

## 16.5 Topographie de l'échantillon tumoral

Nom de l'item	Topographie de l'échantillon tumoral
Numéro de l'item	16.5
Définition	Localisation anatomique de l'échantillon tumoral selon la Classification Internationale des Maladies pour l'Oncologie.
Nomenclature	CIM-O-3

utilisée	
valeur attendue	<p>Pour cette valeur, une précision aussi poussée n'est pas nécessaire pour la localisation des métastases que pour celle des tumeurs primitives.</p> <p>Ainsi, il est suffisant de se limiter au premier niveau de topographie de la classification CIM-O-3.</p> <p>Exemples : :</p> <p>C22    3        Foie</p> <p>C34    3        Poumon</p>
caractère obligatoire ou optionnel	Optionnel

## 16.6 Disponibilité

Nom de l'item	Disponibilité de l'échantillon biologique
Numéro de l'item	16.6
Définition	Indicateur de la disponibilité de l'échantillon pour réutilisation
valeur attendue	Oui/Non
caractère obligatoire ou optionnel	Optionnel
Options	<p>Cette information est importante mais dépend de la manière dont sont gérés les systèmes de conservation des échantillons comme les Centre de ressources biologiques (CRB), les sérothèques, les laboratoires de cyto-pathologie.</p> <p>A titre d'exemple, l'information varie si le patient est vivant ou décédé ou selon la quantité de matériel disponible (« bloc épuisé » par exemple) ou la nature du matériel disponible (bloc, lame blanche, ...).</p> <p>Il n'est pas proposé de conserver l'historique mais il serait nécessaire de qualifier cette information par la date de mise à jour de la disponibilité de l'échantillon</p>

# LES BIOMARQUEURS ET MARQUEURS TUMORAUX

Les biomarqueurs en cancérologie se sont développés de façon très importante. Ils permettent de dépister un cancer, mieux caractériser une tumeur, suivre son évolution, guider le traitement, être un outil prédictif ou de pronostic. Utilisés à de nombreux stades de la maladie, leur importance justifie un chapitre entier de la description d'une tumeur.

Ils sont de nature et de types très différents. Nous proposons d'adopter une convention selon le type d'analyse pour structurer les différents biomarqueurs :

- **Biomarqueurs tissulaires**
  - Plusieurs biomarqueurs tissulaires possibles
  - Rattaché à une tumeur primaire (ou à un évènement tumoral)
  - Evolution possible lors des différentes évolutions tumorales
  - Détecté par l'anatomopathologie (Immuno- histochimie - IHC) ou FISH
  - Ex : Hyper-expression de Her2, RO/RP, translocation de ALK, mutation de EGFR
- **Biomarqueurs moléculaires tumoraux**
  - Plusieurs biomarqueurs moléculaires possibles
  - Rattaché à une tumeur primaire (ou à un évènement tumoral)
  - Evolution possible lors des évolutions tumorales
  - Détecté analyse génomique, par séquençage de l'ADN (exome ou panel de gènes)
  - Biopsie solide / Biopsie liquide
  - Ex : mutation de EGFR, mutation de KRAS, mutation BRAF, réarrangements du gène RET
- **Biomarqueur moléculaires constitutionnels**
  - Un seul biomarqueur par patient
  - Rattaché au patient
  - Pas d'évolution au cours du temps
  - Peuvent être mesurés dans divers échantillons biologiques
  - Prédispositions génétiques à certaines tumeurs (sein, ovaire)
  - Ex : mutation BRCA1
- **Marqueurs tumoraux (ou marqueurs sériques)**
  - Plusieurs marqueurs tumoraux possibles
  - Attaché au patient et à une date
  - suivi du traitement et évolution de la maladie
  - mesuré sur des liquides par des techniques biologiques (par prélèvement sanguin, parfois prélevés dans un épanchement, pleural, ou dans le LCR)
  - Ex : CA15-3, PSA, ACE



### Remarques générales :

A chaque fois que c'est possible, la nomenclature utilisée est LOINC. Néanmoins, comme il n'existe pas un code unique pour un biomarqueur donné, des choix seront nécessaires. Il sera donc nécessaire d'assurer la maintenance d'une liste de réponse et de code LOINC déterminée, d'utiliser une liste de synonyme (Exemple : HER2 / *c-erb B2*).

Il n'existe pas pour les biomarqueurs tissulaires une seule manière de les quantifier ; par exemple, pour KI67 et PDL1, on mesure un pourcentage, pour l'hyper-expression de HER2 en IHC, on attend une classification en +, ++, +++.

Les marqueurs tumoraux sont principalement utilisés pour surveiller l'efficacité des traitements et ne possèdent généralement pas de valeur diagnostique (à l'exception de marqueurs tels que le PSA ou les marqueurs sériques des tumeurs germinales). Certains marqueurs tumoraux ou biomarqueurs tissulaires peuvent être recherchés dans des localisations tumorales particulières. Néanmoins, il n'est pas proposé de lien entre une localisation tumorale, une morphologie particulière et une liste de marqueurs tumoraux ou de biomarqueurs qu'il serait difficile de maintenir. Par contre, lors pour une analyse de qualité des données, il est recommandé de rechercher des liens peu probables ou impossibles.

En revanche, les biomarqueurs tissulaires et moléculaires sont d'une grande valeur informative, soit en raison de leur caractère diagnostique, théranostique ou pronostique. Pour détecter une mutation (Ex : une mutation de EGFR) ou une amplification d'un gène (Ex : gène MET), plusieurs techniques peuvent être employées et ces techniques évoluent au cours du temps (par exemple, passer d'une technique de séquençage à une technique d'IHC). Il est intéressant de coder la technique utilisée.

## 17 – Les biomarqueurs tissulaires

### 17.1. Nom du biomarqueur tissulaire

Nom de l'item	Biomarqueur tissulaire
Numéro de l'item	17.1
Définition	Nom du biomarqueur tissulaire
Nomenclature utilisée	LOINC
caractère obligatoire ou optionnel	Obligatoire
Recommandations	<p>Le référentiel doit permettre un niveau de détail conforme à ce qui est observé dans les soins courants.</p> <p>Ainsi, il est utile de coder les variants ou les exons dans certains cas, notamment ceux concernant les gènes mutés lorsque cette information est disponible cliniquement dans les dossiers médicaux.</p>

	<p>A titre d'exemple, il est utile de récupérer les variants cliniquement pertinents dans la pratique, tels que BRAF V600E ou KRAS G12C.</p> <p>Il est important de présenter les résultats positifs mais également les résultats négatifs ou non contributifs.</p>
--	---

## 17.2. Technique utilisée pour l'analyse du biomarqueur tissulaire

Nom de l'item	Type d'analyse
Numéro de l'item	17.2
Définition	Type d'analyse utilisée pour la détection du biomarqueur moléculaire ou tissulaire
Nomenclature utilisée	LOINC
valeur attendue	IHC FISH CISH
caractère obligatoire ou optionnel	Obligatoire

## 17.3 Résultat du biomarqueur

Nom de l'item	Résultat du biomarqueur
Numéro de l'item	17.3
Définition	Valeur d'expression du biomarqueur
Nomenclature utilisée	-
valeur attendue	Les résultats des biomarqueurs sont séparés en deux variables : - Résultat booléen - Résultat quantitatif
caractère obligatoire ou optionnel	Optionnel
Recommandations	Le résultat attendu est « positif » ou « négatif » ou « non contributif » sur la mutation du gène, l'exon, le variant du gène

## 17.4 Date du résultat

Nom de l'item	Date du résultat
Numéro de l'item	17.4
Définition	Date du résultat pour les biomarqueurs moléculaires et tissulaires
Valeur attendue	YYYY - MM - DD
Règle	<p>Les dates de résultats des biomarqueurs moléculaires ou tissulaires sont essentielles.</p> <p>Lorsque les données sources sont disponibles, c'est la date du résultat qui est utilisée.</p> <p>Cependant, lorsqu'une analyse est effectuée en dehors de l'établissement, les dates sont extraites des dossiers médicaux à l'aide de techniques de traitement du langage naturel (NLP), privilégiant la date du compte rendu.</p> <p>Cette règle s'applique aux biomarqueurs moléculaires et tissulaires où la date du compte rendu est préférée à la date du prélèvement.</p> <p>Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.</p>
justification du choix de la règle (exemple)	<p>Lorsque l'analyse est effectuée en dehors de l'établissement, la date du compte rendu est toujours privilégiée</p> <p>Dans certains cas, l'analyse moléculaire peut être réalisée plusieurs mois, voire années, après la date de la biopsie. La décision du clinicien concernant le début d'une thérapie spécifique est basée sur les résultats de l'analyse, et donc sur cette date.</p> <p>Ainsi, pour aligner la date sur la décision clinique, la date du compte rendu est utilisée comme référence.</p>
caractère obligatoire ou facultatif	Facultatif

## 17. Les biomarqueurs moléculaires tumoraux

### 17.5 Nom du biomarqueur moléculaire

Nom de l'item	Biomarqueur moléculaire
Numéro de l'item	17.15
Définition	Nom du biomarqueur moléculaire
Nomenclature	LOINC

utilisée	
caractère obligatoire ou optionnel	Obligatoire
Recommandations	<p>Le référentiel doit permettre un niveau de détail conforme à ce qui est observé dans les soins courants.</p> <p>Ainsi, il est utile de coder les variants ou les exons dans certains cas, notamment ceux concernant les gènes mutés lorsque cette information est disponible cliniquement dans les dossiers médicaux.</p> <p>A titre d'exemple, il est utile de récupérer les variants cliniquement pertinents dans la pratique, tels que BRAF V600E ou KRAS G12C.</p> <p>Il est important de présenter les résultats positifs mais également les résultats négatifs ou non contributifs.</p>

## 17.6. Technique utilisée pour l'analyse du biomarqueur moléculaire

Nom de l'item	Type d'analyse
Numéro de l'item	17.6
Définition	Type d'analyse utilisée pour la détection du biomarqueur moléculaire ou tissulaire
Nomenclature utilisée	LOINC
valeur attendue	IHC FISH Etc.
caractère obligatoire ou optionnel	Obligatoire

## 17.7 Résultat du biomarqueur

Nom de l'item	Résultat du biomarqueur
Numéro de l'item	17.7
Définition	Valeur d'expression du biomarqueur
valeur attendue	<p>Les résultats des biomarqueurs sont séparés en deux variables :</p> <ul style="list-style-type: none"> <li>- Résultat booléen</li> <li>- Résultat quantitatif</li> </ul>
caractère obligatoire ou optionnel	Optionnel

## 17.8 Date du résultat

Nom de l'item	Date du résultat
Numéro de l'item	17.8
Définition	Date du résultat pour les biomarqueurs moléculaires et tissulaires
Valeur attendue	YYYY - MM - DD
Règle	<p>Les dates de résultats des biomarqueurs moléculaires ou tissulaires sont essentielles.</p> <p>Lorsque les données sources sont disponibles, c'est la date du résultat qui est utilisée.</p> <p>Cependant, lorsqu'une analyse est effectuée en dehors de l'établissement, les dates sont extraites des dossiers médicaux à l'aide de techniques de traitement du langage naturel (NLP), privilégiant la date du compte rendu.</p> <p>Cette règle s'applique aux biomarqueurs moléculaires et tissulaires où la date du compte rendu est préférée à la date du prélèvement.</p> <p>Si une valeur est non disponible sur le jour (DD), le mois (MM) ou l'année (YYYY). Le champ correspondant sera laissé sans valeur.</p>
justification du choix de la règle (exemple)	<p>Lorsque l'analyse est effectuée en dehors de l'établissement, la date du compte rendu est toujours privilégiée</p> <p>Dans certains cas, l'analyse moléculaire peut être réalisée plusieurs mois, voire années, après la date de la biopsie. La décision du clinicien concernant le début d'une thérapie spécifique est basée sur les résultats de l'analyse, et donc sur cette date.</p> <p>Ainsi, pour aligner la date sur la décision clinique, la date du compte rendu est utilisée comme référence.</p>
caractère obligatoire ou facultatif	Facultatif

## 17. Les biomarqueurs moléculaires constitutionnels

### 17.9 Nom du biomarqueur constitutionnel

Nom de l'item	Nom du biomarqueur constitutionnel
Numéro de l'item	17.9
Définition	Un biomarqueur constitutionnel, également appelé biomarqueur génétique, fait référence à une caractéristique biologique intrinsèque,

	présente dès la naissance, généralement héritée, qui est associée à un risque accru de développer un cancer.
Nomenclature utilisée	LOINC
valeur attendue	Chaine de caractères  Les principaux biomarqueurs constitutionnels sont les suivants (cette liste n'est pas exhaustive) : BRCA1 BRCA2 Etc. (liste à compléter)

## 17.10 Résultat de la mise en évidence biomarqueur constitutionnel

Nom de l'item	Résultat de l'analyse du biomarqueur constitutionnel
Numéro de l'item	17.10
Définition	Valeur d'expression du biomarqueur constitutionnel
valeur attendue	Mutation du gène Pas de mutation du gène Résultat non contributif
caractère obligatoire ou optionnel	Obligatoire

## 17. Les marqueurs tumoraux

### 17.11 Nom du marqueur tumoral

Nom de l'item	Nom du marqueur tumoral
Numéro de l'item	17.11
Définition	Les marqueurs tumoraux sont des substances circulantes, souvent des protéines ou des hormones, qui sont produites par les cellules cancéreuses (ou parfois par l'organisme en réponse à la croissance d'un cancer). Ces substances sont dosées dans le sang, le LCR, la plèvre.
Nomenclature utilisée	LOINC
valeur attendue	Chaine de caractères

	Les principaux marqueurs tumoraux sont : PSA ACE (Antigène carcino-embryonnaire) CA 125 CA 19-9 CA 15-3 SCC (Squamous Cell Carcinoma Antigen) HCG (Hormone Chorionique Gonadotrope) α-foeto protéine (AFP) Cyfra 21 Thyrocalcitonine NSE
--	---

### 17.12 Résultat de l'analyse du marqueur tumoral

Nom de l'item	Résultat de l'analyse du marqueur tumoral
Numéro de l'item	17.12
Définition	Valeur d'expression du marqueur tumoral
valeur attendue	Valeur numérique du marqueur tumoral. Par exemple si le PSA est à 12 on prend la valeur « 12 » sans référence à une valeur normale et sans spécifiée l'unité utilisée
caractère obligatoire ou optionnel	Obligatoire

### 17.13 Nature du prélèvement (marqueur tumoral)

Nom de l'item	Nature du prélèvement pour le dosage du marqueur tumoral
Numéro de l'item	17.13
Définition	Nature du prélèvement pour le dosage du marqueur tumoral
valeur attendue	Sérum, LCR, liquide pleural
caractère obligatoire ou optionnel	Optionnel

### 17.14 Date du résultat (marqueur tumoral)

Nom de l'item	Date du résultat
Numéro de l'item	17.14
Définition	Date du résultat pour les marqueurs tumoraux

valeur attendue	YYYY - MM - DD
Règle	<p>Les dates des résultats des marqueurs tumoraux sont essentielles car elles permettent un suivi de l'évolution dans le temps.</p> <p>Lorsque la source biologie est disponible, la date de prélèvement est privilégiée.</p> <p>Cependant, lorsqu'une analyse est effectuée en dehors de l'établissement, les dates sont extraites des dossiers médicaux à l'aide de techniques de traitement du langage naturel (NLP), privilégiant la date du compte rendu.</p>
caractère obligatoire ou optionnel	Obligatoire



# ANNEXE : LES LIGNES DE TRAITEMENT

## Définition

Une ligne de traitement, également appelée "ligne de chimiothérapie" ou "protocole de traitement", désigne un ensemble de médicaments administrés à un patient atteint de cancer pendant une période donnée.

Une ligne de traitement contient plusieurs cycles et un cycle correspondant à un médicament ou une association de médicaments, administré à intervalle régulier.

La notion de ligne de traitement est difficile à définir et il n'existe pas de consensus absolu. Ce texte a pour but de spécifier au mieux la façon de définir une ligne dans le modèle OSIRIS RWD et afin que tous alimentent l'item « ligne de traitement » selon les mêmes règles et recommandations.

Les lignes de traitement concernent les chimiothérapies cytotoxiques, les traitements par voie orale (hormonothérapie, thérapie ciblée orale type TKI), les anticorps monoclonaux, l'immunothérapie.

## Principes généraux

- La notion de ligne de traitement ne s'applique que pour les patients en phase métastatique.
- On repère de façon automatique les changements de molécules administrées
- La source structurée est privilégiée sur les documents non structurés

## Recommandations

On réserve la notion de ligne de traitement aux patients en phase métastatique. Ainsi, on compte le nombre de lignes à partir de la première administration en phase métastatique.

- **Première ligne de traitement** : Il s'agit du premier ensemble de médicaments administrés à un patient dès le diagnostic de cancer métastatique, sans tenir compte des traitements adjuvants.
- **Deuxième ligne de traitement** : Si la première ligne de traitement n'est pas suffisamment efficace ou si la maladie progresse ou en cas de toxicité, il est proposé une deuxième ligne de traitement.
- Etc.

*Exemple : si un patient est traité par chimiothérapie adjuvante, et après une phase de rémission et de surveillance, il apparaît une maladie métastatique, le premier protocole de chimiothérapie est un traitement de 1<sup>ère</sup> ligne.*

La notion de ligne de traitement ne s'applique pas en situation adjuvante et on ne calcule pas de ligne dans ce cas.

## Hormonothérapies et thérapies ciblées

Lorsqu'un patient est sous hormonothérapie ou thérapie ciblée orale, on compte les lignes pour l'hormonothérapie ou les thérapies ciblées orales de la même façon.

*Exemple 1 : une patiente sous tamoxifène en 1<sup>ère</sup> ligne progresse ; on passe une anti-aromatase ; il s'agit d'un traitement de 2<sup>ème</sup> ligne.*

*Exemple 2 : dans la suite, si on constate une progression, on passe à une chimiothérapie, il s'agit de la 3<sup>ème</sup> ligne en situation métastatique.*

## Cas particulier des cancers hématologiques

On compte comme les tumeurs solides ; on ne parle pas de récurrence mais de progression. Les mêmes règles s'appliquent que pour les tumeurs solides métastatiques.

## Cas particulier des traitements de maintenance

Dans certains cas, après une rémission ou une stabilisation de la maladie, un patient peut être placé sous traitement de maintenance, qui consiste à poursuivre un des médicaments administrés dans la ligne précédente pour maintenir la rémission ou contrôler la maladie sur une période prolongée. Certains appellent ce traitement, chimiothérapie de continuation.

*Exemple : le traitement standard des adénocarcinomes pulmonaires associe le cisplatine et le pémétréxed (4 cycles) suivi de l'utilisation du pémétréxed en continuation tant que le traitement est efficace, c'est-à-dire tant qu'il n'y a pas de progression de la maladie, et tant que le traitement est bien supporté. La poursuite du pémétréxed ne constitue pas une nouvelle ligne.*

Chimiothérapie comportant A + B puis A seul : pas de nouvelle ligne

## Cas particulier du switch de certaines molécules en raison d'une toxicité

Dans certains cas, un des médicaments a une toxicité qui oblige à l'arrêt ; il peut alors être remplacé par un médicament de la même classe mais avec un profil de toxicité différent ou par un médicament également efficace mais appartenant à une autre classe avec un mode d'action différent.

*Exemple : le traitement standard des carcinomes épidermoïdes bronchiques associe le cisplatine et la vinorelbine (4 cycles). Lorsqu'il apparaît une toxicité rénale, le cisplatine peut être remplacé par le carboplatine. Lorsqu'il apparaît une toxicité digestive à type d'occlusion, la vinorelbine peut être remplacée par la gemcitabine.*

Un débat agite les oncologues médicaux ; pour certains, il s'agit bien d'une nouvelle ligne, pour d'autres, une simple adaptation. On propose par convention et pour permettre une automatisation du processus de convenir qu'il s'agit d'une nouvelle ligne.

Chimiothérapie comportant A + B puis A + C : nouvelle ligne

## Cas de l'arrêt d'un des médicaments lors d'une association

Dans certains cas, pour des protocoles à 3 médicaments, un des médicaments a une toxicité qui oblige à l'arrêt ; il n'est pas remplacé mais définitivement arrêté.

Chimiothérapie comportant A + B + C puis A + B : pas de nouvelle ligne

### Cas d'une diminution de doses

Lors d'un traitement, quel que soit la molécule administrée, il est parfois diminuer les doses d'un ou des médicaments administrés en association. Il ne s'agit pas d'une nouvelle ligne

### Cas de l'augmentation de l'intervalle de temps entre les cycles

Si les médicaments administrés ne changent pas, il ne s'agit pas d'une nouvelle ligne, quel que soit l'intervalle de temps entre les cycles y compris si le délai est très long.

### Cas des protocoles qui alternent des associations différentes

Le nom du protocole n'est pas utilisable car le paramétrage dépend du centre et il est impossible, dans un protocole, de détecter des changements de ligne qui n'en sont pas. Ceci peut constituer une limite des règles et des conventions mais cette situation est assez rare ; on la retrouve dans certains protocoles en pédiatrie.

### Sources utilisées

Trois sources peuvent être utilisées :

- Les logiciels de prescription type CHIMIO® avec une information structurée
- Les comptes rendus médicaux de l'hôpital avec une information en texte libre
- Les documents récupérés de l'extérieur de l'hôpital et scannés dans le dossier (texte libre).

### La source des données structurées

La meilleure source de données est celle issue des outils de prescription de chimiothérapie comme CHIMIO® ou d'autres logiciels de prescription de chimiothérapie.

Les données récupérées sont les traitements administrés par molécule avec les dates d'administration (et non les traitements simplement prescrits)

La source est structurée pour les traitements systémiques à base de molécules anti-cancer avec des informations précises :

- Molécules en DCI (docetaxel) ou spécialités pharmaceutiques (TAXOTERE®), en ne tenant compte que des doses réellement administrées (et non des doses prévues)
- Date d'administration

La source CHIMIO dispose uniquement des molécules injectées (et/ou qui nécessitent une reconstitution

- Concerne les patients traités dans le centre
- Les traitement IV uniquement (sauf exception)
- La chimiothérapie conventionnelle, l'immunothérapie, les anticorps monoclonaux
- Il n'est pas tenu compte des : « traitements associés » (acide folinique, corticoïdes...)

### Les comptes rendus médicaux

Certains traitements ne se trouvent pas dans la source structurée :

- Patients non traités dans l'établissement

- Traitements oraux : hormonothérapie, chimiothérapie orale, thérapie ciblée

Dans ce cas, il est nécessaire de reconstituer le traitement pour structurer l'information à partir des dossiers médicaux.

Plusieurs sources d'informations sont susceptibles d'occasionner des doublons ; il est proposé une fusion des cycles de la source chimio avec les comptes rendus

## PARTICIPANTS AU GROUPE DE TRAVAIL

1. Alain Livartowski (Médecin Oncologue, Unicancer, coordonnateur du groupe)
2. Olivier Carsenti (PSCC, coordinateur-adjoint du groupe)
3. Emmanuel Desandes (Institut de cancérologie de Lorraine)
4. Toufik Belguendouz Bendahmane (Médecin Oncologue)
5. Aurélien Jobard (Chef de projet Osiris, Inca)
6. Alexandre Yazigi (Chirurgien Oncologue CEO, Gimli)
7. Ali Bellamine (Médecin Santé Publique Data scientist, Elsan)
8. Aurélie Bertaut (Médecin - Centre Georges Francois Leclerc)
9. Mathieu Robain (Directeur scientifique, Unicancer)
10. Vincent Le Texier (Chef de Projet Données de santé, Fondation Synergie Lyon Cancer)
11. Julie Perlberg Samson (Médecin DIM, Institut Curie)
12. Francois Lion (Gustave Roussy)
13. Laetitia Chanas, (Direction des data, Institut curie)
14. Marc Deloger (Responsable service bio-informatique, Gustave Roussy)
15. Myriam Kossai (Médecin anatomo-pathologiste, CLCC Clermont Unicancer)
16. Elodie Pleynet (Centre Lyon Bérard - Unicancer)
17. Yec'han Laizet (Institut Bergonié)
18. Guillaume Danton (Data Scientist, Centre Eugène Marquis)
19. Julien Huyard (Data Manager, Unicancer)
20. Samantha Padeloup (Elsan)
21. Raphaël Sigogne (Responsable Data, Paris Saclay Cancer Cluster)
22. Hugo Crochet (Directeur des Systèmes d'Information, Centre Léon Bérard)
23. Pierre Heudel (Médecin Oncologue, Centre Lyon Bérard)
24. Francois Bocquet (Institut de Cancérologie de l'Ouest Nantes-Angers)

Remerciements à Coralie Courtinard, Emilie Bensadon, Catherine Eng pour leurs relectures attentives