

Sources d'information pour les registres du cancer basés sur la population

L'utilisation de multiples sources d'information sur les cas de cancer survenant dans la population cible représente un élément essentiel des registres de population. Cela permet d'identifier le plus de cas possible parmi ceux qui sont diagnostiqués chez les habitants de la zone géographique couverte par le registre. Cela ne pose pas de problème si l'on reçoit des informations sur les mêmes cas de cancer à partir de plusieurs sources (comme décrit dans le chapitre 5, cela permet d'évaluer la capacité du registre de population à trouver les cas). Les procédures utilisées par les registres permettent d'identifier le même cas de cancer à partir de différentes sources (tout en évitant les doublons) ; cette disposition est intégrée dans le logiciel CanReg5 (voir l'Annexe 1).

1. Sources d'information sur les cas de cancer

Les sources de données peuvent être groupées en trois grandes catégories, dont chacune sera discutée plus loin :

- hôpitaux
- laboratoires
- certificats de décès.

1.1 Sources hospitalières

Le registre doit tenter d'identifier tous les cas de cancer diagnostiqués ou traités dans les hôpitaux ou les cliniques situés dans la zone couverte. Les établissements concernés vont varier en fonction de leur situation géographique, mais il est important de tous les identifier et tous les énumérer, et de déterminer le nombre probable de patients cancéreux (ainsi que le type de cancer) vus dans

chacun d'eux. Quand ils existent, les centres spécialisés dans le traitement du cancer (oncologie médicale/chirurgicale, radiothérapie) apportent au registre une contribution essentielle. Souvent, ces services possèdent un registre des cas diagnostiqués, traités, ou suivis.

La plupart des services hospitaliers voient des patients atteints de cancer, même si la proportion des cas de cancer varie en fonction de leur spécialisation. Pour la recherche des cas, le registre va utiliser en premier lieu, lorsqu'ils existent, les systèmes d'information hospitaliers, d'où l'on peut extraire les informations sur les patients et sur leur diagnostic. Même si le système d'information hospitalier n'est pas informatisé, le service de gestion des dossiers médicaux peut maintenir à jour manuellement le répertoire des sorties de l'hôpital, dans lequel on peut effectuer un tri

par diagnostic. Quand il n'existe pas de système centralisé de gestion des informations, le travail du registre est plus laborieux et peut nécessiter des visites auprès de chacun des services cliniques.

Les hôpitaux privés et les cliniques ont tendance à être de plus petite taille que les grands hôpitaux publics, et ne disposent pas toujours d'installations spécialisées pour le traitement du cancer. Néanmoins, il peut être important de les inclure dans les sources de données, si l'identification des patients cancéreux parmi leur clientèle est relativement facile. Il se peut que les détenteurs des données soulèvent des problèmes de confidentialité (réels ou imaginaires) relatifs à leur collaboration avec le registre du cancer.

Les hospices et les services de soins palliatifs représentent des sources d'information très importantes. La grande majorité de leurs clients sont des patients cancéreux, la documentation du diagnostic est généralement bonne, et le suivi jusqu'à la mort est la norme (c'est en effet leur raison d'être).

1.2 Services de laboratoire

Les laboratoires d'anatomopathologie représentent une source majeure et essentielle de données. Pour la plupart des patients cancéreux, le diagnostic définitif est basé sur l'histologie (bien que la proportion des cas pour lesquels la tumeur est examinée par un anatomopathologiste dépend du site d'origine et du type de cancer). Les laboratoires d'anatomopathologie consignent toujours leurs rapports dans un registre – souvent une base de données informatisée, mais même quand les registres sont sous forme papier, il est facile de scanner les diagnostics de cancer. Mais en ce qui concerne les informations sur les patients cancéreux, les laboratoires dépendent souvent

des formulaires de demande qui leur sont adressés avec le prélèvement. Ces formulaires peuvent contenir des informations insuffisantes ou ne pas avoir été remplis entièrement – spécialement en ce qui concerne le lieu de résidence du patient. Cette information est essentielle pour les registres de population, et il est nécessaire de faire un effort particulier pour la trouver pour les cas identifiés par le biais des laboratoires.

Les autres services de laboratoire sont de moins bonnes sources d'information, même si l'examen des échantillons de moelle osseuse (et donc le diagnostic d'hémopathie maligne) est généralement effectué par les hématologues cliniciens (plutôt que par les anatomopathologistes). Dans les services d'imagerie médicale, seules l'imagerie par résonance magnétique (IRM) et la tomodensitométrie (TDM) voient suffisamment de cas de cancer pour que l'on puisse les considérer comme sources de données. Leur utilité dépend de la facilité avec laquelle les cas de cancer peuvent être identifiés dans les listes de patients examinés.

1.3 Certificats de décès

Les informations sur les personnes décédant de cancer (ou avec un cancer) constituent de très importantes sources de données pour les registres. Ces informations peuvent provenir des registres d'état civil (où la « cause de décès » est mentionnée par un médecin sur le certificat de décès), même si ce processus est incomplet (les décès ne donnant pas tous lieu à un certificat). Dans les registres d'état civil des PRFI, l'attribution et le codage des causes de décès posent souvent des problèmes. Dans beaucoup de pays à revenu faible, l'enregistrement des décès se limite à ceux qui surviennent à l'hôpital (pas de certificat délivré par un médecin pour les décès se produi-

sant à la maison) ; même avec leurs limites, ces données doivent être exploitées par le registre.

L'identification des personnes décédant d'un cancer (ou avec un cancer) répond à trois besoins du registre :

- Elle permet l'identification de cas de cancer qui ont été « manqués » par le système de collecte des données.
- Elle permet d'enregistrer le décès de patients cancéreux figurant déjà dans le registre (utile pour le calcul de la survie).
- La connaissance du nombre de cas notifiés pour la première fois grâce au registre des décès permet d'évaluer l'exhaustivité du registre du cancer.

2. Collecte des données

Traditionnellement, on distingue deux types de méthode de recueil des données : la collecte « passive » (reposant sur le personnel médical qui renseigne les formulaires de notification et les adresse au registre) et la collecte « active », selon laquelle le personnel du registre du cancer rend visite aux diverses sources de données pour identifier et résumer les informations pertinentes. L'enregistrement ne fonctionne pas s'il repose entièrement sur la diligence et la bonne volonté des personnes étrangères à l'équipe du registre pour le recueil et la synthèse des informations sur les cas de cancer. La plupart des registres associent néanmoins ces deux méthodes, et si la recherche active des cas reste la norme, le développement de l'informatisation des données médicales permet une certaine utilisation des bases de données électroniques pour la recherche des cas.

Avec l'augmentation du nombre de sources de données informatisées disponibles, les registres du cancer sont parfois soumis à des

Tableau 4.1. Données de base pour les registres du cancer

Variable	Commentaires
L'individu	
<i>Identification^a</i>	
Nom	Selon l'usage local
Sexe	
Date de naissance ou âge	Estimés si non connus
<i>Données démographiques</i>	
Adresse	Résidence habituelle
Groupe ethnique ^b	Quand la population est composée de deux ou plusieurs groupes
La tumeur	
Date d'incidence	
Base de diagnostic	
Topographie (siège)	Tumeur primitive
Morphologie (histologie)	
Comportement	
Source d'information	Exemple : numéro d'enregistrement à l'hôpital, nom du médecin

^a Le minimum recueilli est ce qui assure que, si le même individu est notifié plusieurs fois au registre, il sera identifié comme étant la même personne. Ce peut être également un numéro d'identification personnel unique.

^b Le groupe ethnique est inclus ici en raison de son importance pour un certain nombre de registres, particulièrement dans les pays en développement. Source : MacLennan (1991, 1996).

pressions les poussant à abandonner leurs modes traditionnels de fonctionnement. Il est évident que sur le long terme, les registres vont devoir développer une stratégie leur permettant de passer des sources de données papier aux sources de données informatisées, mais il est faux de croire que les données des registres du cancer peuvent être dérivées automatiquement des données des systèmes d'information médicale. Indépendamment des sources d'information et des méthodes de recueil des données utilisées, la production de données d'incidence de haute qualité nécessite un personnel qualifié. Dans certains PRFI, le temps dont dispose le personnel des registres ne permet que le traitement de routine des données et la production des données d'incidence. L'utilisation des données provenant des systèmes d'information médicale pourrait néanmoins permettre au personnel du registre de passer moins de temps à la saisie des données et d'en consacrer plus

au contrôle de qualité, à l'analyse des données et éventuellement à la recherche.

3. Variables collectées par les registres

Pour chaque cas de cancer, les registres ont à rentrer les données pour tout un ensemble de variables. On a toujours tendance, lors de la planification du registre, à vouloir inclure trop de variables. Il ne faut pas oublier que les informations collectées proviennent de sources secondaires (comptes rendus cliniques, rapports d'anatomopathologie, extraits des registres des sorties des hôpitaux, certificats de décès) et NON PAS des patients eux-mêmes. Aussi convient-il d'éviter d'inclure les données qui ne sont pas systématiquement disponibles à partir de ces sources. Cela s'applique en particulier aux informations qui ne peuvent s'obtenir de façon fiable qu'en interrogeant le patient (facteurs de risque comme le tabac et la consommation

d'alcool, l'alimentation, etc.), et à celles qui risquent de n'être enregistrées que dans un sous-groupe particulier de patients (et non pas un sous-groupe aléatoire) comme la profession ou le statut par rapport au VIH. En règle générale, les informations qui ne peuvent pas être collectées de façon fiable dans 80–90% des cas ne doivent pas figurer dans le registre. Il vaut mieux éviter également d'inclure les informations peu pertinentes (comme le statut marital), même si elles sont faciles à recueillir. Le manuel *Enregistrement des cancers : Principes et méthodes* propose un ensemble de 10–11 variables essentielles, constituant l'*information de base* à collecter (Tableau 4.1) ; il est vrai qu'aucun registre du cancer ne peut fonctionner avec moins de données, ce qui fait que l'on peut les considérer comme le **minimum de données à collecter**.

La liste raisonnable des **variables essentielles** est toutefois plus substantielle. Le Tableau 4.2 se fonde sur les recommandations du

Tableau 4.2. Variables essentielles pour les registres du cancer

Variable	Commentaires
L'individu	
Identification	Dans certains pays, numéro d'identification personnel unique, dans d'autres, nom complet combiné avec la date de naissance et le sexe
Date de naissance	Jour, mois et année (jj/mm/aaaa)
Sexe	Masculin (M) ou féminin (F)
Groupe ethnique	En fonction de la situation locale
Adresse avec le code postal (et le numéro de téléphone)	Nécessaire pour l'identification et pour les analyses géographiques
La tumeur	
Date d'incidence	Information prioritaire, comme souligné dans les recommandations de l'ENCR.
Localisation anatomique (topographique) de la tumeur primitive	Au minimum selon la classification CIM-O.
Latéralité	A enregistrer pour tous les organes allant de pair, mais au moins pour les organes suivants : sein, œil, ovaire, testicule et rein (mais en observant les règles établies pour les tumeurs primitives multiples).
Type histologique de la tumeur primitive	Au minimum selon la classification CIM-O.
Comportement	Au minimum selon la classification CIM-O.
Fondements du diagnostic	Il est recommandé d'utiliser les fondements les plus vraisemblables sur lesquels peut s'appuyer le diagnostic. Il est possible d'enregistrer toutes les méthodes pertinentes. Les codes utilisés pour les fondements du diagnostic de cancer doivent utiliser la classification CIM-O.
Stade – (TNM condensée)	Le stade est nécessaire pour les études internationales et pour les cliniciens. Il est recommandé d'utiliser la classification TNM condensée de l'ENCR.
Traitement initial (initié dans les 4 mois suivant la date d'incidence) [Les registres devraient disposer d'un manuel clair sur les données à inclure pour tous les items relatifs aux traitements.]	Au minimum, les registres devraient pouvoir présenter les modalités de traitement utilisées sous forme de réponses oui/non.
<i>Chirurgie</i>	Tout acte chirurgical à visée curative ou palliative
<i>Radiothérapie</i>	Toute radiothérapie à visée curative ou palliative
<i>Chimiothérapie</i>	Toute chimiothérapie anticancéreuse à visée curative ou palliative
<i>Thérapie endocrine (hormones)</i>	Thérapie exogène (médicament)
Sources d'information	
Sources d'information	Il est important d'enregistrer TOUTES les sources d'information (hôpital, établissement) pour chaque modalité de diagnostic et de traitement, de façon à pouvoir effectuer un contrôle de qualité, ou collecter des informations supplémentaires. La date et le numéro de l'hôpital ou du laboratoire sont enregistrés à chaque fois.
Suivi/mise à jour	
Date de la dernière mise à jour	Nécessaire pour étudier le suivi (jj/mm/aaaa)
Statut vital (à la date de la mise à jour)	Il peut être intéressant d'indiquer si connu ou supposé (par exemple, par recoupement avec les certificats de décès) (jj/mm/aaaa)
Date du décès	Nécessaire pour étudier la survie et le suivi (jj/mm/aaaa)

CIM-O, Classification internationale des maladies pour l'oncologie ; ENCR, Réseau européen des registres du cancer ; TNM, tumeur – ganglions – métastases.

Source : *Recommendations for a Standard Dataset for the European Network of Cancer Registries* [Recommandations relatives aux données normalisées pour le Réseau européen des registres du cancer – en anglais uniquement] (<http://www.enccr.eu/images/docs/recommendations/recommendations.pdf>).

Réseau européen des registres du cancer (ENCR) (<http://www.enccr.eu/images/docs/recommendations/recommendations.pdf>).

Il existe de nombreuses **variables optionnelles** qui peuvent

également être incluses en fonction des intérêts locaux particuliers, en tenant compte des considérations sur la disponibilité des informations dans les sources de données, comme décrit ci-dessus.

4. Codage

Plusieurs des variables figurant dans la liste doivent être codées pour en faciliter l'analyse. Pour un certain nombre de variables, il existe des

systèmes de codage internationaux standardisés, et les registres du cancer doivent les utiliser pour que les résultats des différents registres soient comparables.

Les systèmes de codage les plus importants concernent le codage de la tumeur (site anatomique, histologie, comportement, base de diagnostic) selon la *Classification internationale des maladies pour l'oncologie* (CIM-O) et le codage de leur extension anatomique (stade) selon le système tumeur – ganglions – métastases (TNM).

On aura besoin, en outre, de systèmes de codage locaux pour :

- le lieu de résidence du patient
- le groupe ethnique (si enregistré)
- les sources d'information.

4.1 Classification des cancers – Classification internationale des maladies pour l'oncologie

La classification CIM-O, maintenant à sa troisième édition, est utilisée depuis plus de 35 ans ; c'est l'outil de référence pour le codage du diagnostic dans les registres du cancer.

La CIM-O est une classification multiaxiale de la topographie, de la morphologie, du comportement et du stade des tumeurs (qui fournit en outre des codes normalisés pour la base de diagnostic).

Le code topographique décrit le **site d'origine** (localisation anatomique) du cancer (le *site primitif* de la tumeur, et non pas la localisation de métastases) et utilise les catégories à trois et quatre caractères figurant dans le chapitre II « Classification des tumeurs » de la *Classification statistique internationale des maladies et des problèmes de santé connexes*, dixième révision (CIM-10) (sauf pour les catégories relatives aux tumeurs secondaires et à des types morphologiques particuliers de tumeurs). La CIM-O offre donc plus de détails sur la localisation des

tumeurs que la CIM-10. A la différence de la CIM-10, la CIM-O inclut les données topographiques des tumeurs hématopoïétiques et réticuloendothéliales (ainsi que d'autres cancers qui, dans la CIM-10, sont définis par l'histologie, comme le sarcome de Kaposi, le mélanome et les sarcomes des tissus mous et des os).

L'axe morphologique fournit des codes à cinq chiffres allant de M-8000/0 à M-9989/3. Les quatre premiers chiffres indiquent le terme histologique spécifique. Le cinquième chiffre, après la barre oblique (/) est le code du comportement tumoral, qui indique si la tumeur est maligne, bénigne, in situ ou de comportement incertain (bénigne ou maligne).

Un code séparé, à un chiffre, est également proposé pour indiquer le grade histopathologique (ou degré de différenciation).

La troisième édition de la *Classification internationale des maladies pour l'oncologie* (CIM-O-3) comprend cinq grandes parties. La première fournit un mode d'emploi général des systèmes de codage et donne les règles pour leur utilisation dans les registres du cancer et les laboratoires d'anatomopathologie. La deuxième partie comprend la liste numérique des codes topographiques, et la troisième, la liste numérique des codes morphologiques. L'index alphabétique combiné proposé dans la quatrième partie donne à la fois les codes topographiques et les codes morphologiques et inclut une sélection de lésions et pathologies pseudo-tumorales. La cinquième partie décrit les différences dans les codes morphologiques entre la deuxième et la troisième édition.

La CIM-O utilise, autant que possible, la nomenclature publiée dans les livres de la collection *Classification des tumeurs* de l'Organisation mondiale de la Santé (« *Blue*

Books » de l'OMS). Au cours des révisions et avec l'introduction de nouveaux termes de morphologie, de nouveaux codes sont mis à jour ou ajoutés à la CIM-O, dont la quatrième édition est en cours de préparation.

La CIM-O est publiée dans de nombreuses langues (allemand, anglais, chinois, coréen, espagnol, finnois, flamand/néerlandais, français [http://whqlibdoc.who.int/publications/2008/9789242545340_fre.pdf], japonais, portugais, roumain, tchèque et turc). On peut se procurer auprès de l'OMS (<http://www.who.int/classifications/icd/adaptations/oncology/en/#>) ou de l'Association internationale des registres du cancer (IACR), pour les membres de cette organisation. Un dossier peut être téléchargé à partir du site de l'OMS (<http://apps.who.int/classifications/apps/icd/ClassificationDownloadNR/login.aspx?ReturnUrl=%2fclassifications%2fapps%2ficd%2fClassificationDownload%2fDLArea%2fDownload.aspx>).

Le pack CIRC–IACR d'outils informatiques pour les registres du cancer (*IARCcrgTools*) comprend des programmes de traitement par lots pour effectuer la conversion de CIM-O vers CIM-10. Les programmes de conversion et de vérification ne peuvent traiter que les fichiers à champs de longueur fixe, même si une option « Transfert de fichiers » permet la conversion d'un fichier texte à champ délimité en fichier à champ de longueur fixe. Le pack IARCcrgTools est disponible à partir du site internet de l'IACR ou du CIRC (http://www.iacr.com.fr/iacr_iarcrgtools.htm).

4.2 Système de codage TNM

La classification TNM de l'Union internationale contre le cancer (UICC) est la norme mondialement acceptée pour décrire le stade des

cancers. Il s'agit d'un système fondé sur l'anatomie qui permet de spécifier l'étendue de la tumeur primitive et son extension régionale aux ganglions lymphatiques voisins, ainsi que l'absence ou la présence de métastases.

Chacune des lettres de TNM correspond à une catégorie :

- La catégorie T décrit la localisation de la tumeur primitive.
- La catégorie N décrit la propagation du cancer aux ganglions lymphatiques régionaux.
- La catégorie M décrit la présence (ou l'absence) de métastases à des sites distants.

La définition du stade du cancer n'est pas seulement importante dans la pratique clinique ; elle fournit aussi des informations vitales pour les responsables de l'élaboration ou de la mise en œuvre des programmes de prévention et de lutte contre le cancer. Il est donc important d'inclure la classification TNM dans les registres du cancer.

La classification TNM est régulièrement mise à jour, la dernière étant la septième édition (2009).

Le site internet de l'UICC donne des explications sur l'utilisation du système de classification TNM et sur la façon dont on peut se procurer les manuels de codage (<http://www.uicc.org/resources/tnm>).

Dans les PRFI, les équipes des registres peuvent avoir des difficultés à rentrer le code TNM complet à partir des dossiers médicaux si les données sur l'extension de la maladie ne sont pas explicitement enregistrées par les cliniciens ou les anatomopathologistes. C'est pourquoi une version simplifiée a été créée par le Réseau européen des registres du cancer, la classification TNM condensée, disponible en anglais et en français ([http://afcrn.org/attachments/article/116/extentofdisease\(condensed%20TNM\)_french%20\(1\).pdf](http://afcrn.org/attachments/article/116/extentofdisease(condensed%20TNM)_french%20(1).pdf)).

Cette version permet la codification de T et/ou N et/ou M quand ils ne sont pas explicitement codés dans les comptes rendus cliniques et/ou anatomopathologiques. Le registre du cancer doit alors essayer de coder l'extension de la maladie selon la classification TNM condensée :

T :	L	A	X
N :	0	+	X
M :	0	+	X

(**A**, avancée ; **L**, localisée ; **X**, inconnu), où T et N proviennent, si possible, du compte rendu anatomopathologique ou, en son absence, du dossier clinique (endoscopie, radiographie, etc.). M est défini en utilisant la meilleure information disponible, qu'elle soit d'ordre clinique, paraclinique ou anatomopathologique. Pour M, un diagnostic clinique est suffisant pour classer le cancer en M+ en l'absence de preuve histologique de métastases.

La classification TNM et la classification TNM condensée permettent toutes deux d'exprimer l'extension tumorale selon le système de classification numérique des stades :

- I Cancer localisé (TL/N0/M0)
- II Cancer avec envahissement local (TA/N0/M0)
- III Cancer avec envahissement régional (tous les T/N+/M0)
- IV Cancer avancé (métastatique) (tous T/tous les N/M+).

4.3 Systèmes de codage locaux

4.3.1 Lieu de résidence

Les codes du « lieu de résidence » doivent correspondre aux subdivisions nationales, telles qu'elles figurent dans les publications des statistiques nationales, et pour lesquelles on dispose d'informations sur la taille et la composition de la population. Il doit être possible de

développer un système de codage hiérarchique, avec plusieurs niveaux de subdivisions (région, province, district, quartier, etc.).

4.3.2 Groupe ethnique

Les codes pour les « groupes ethniques » doivent correspondre, dans la mesure du possible, aux catégories officielles reconnues dans la publication des statistiques nationales, surtout s'il existe des informations disponibles sur la taille et la composition de la population par groupe ethnique.

4.3.3 Sources d'information

Les codes pour les « sources d'information » seront presque toujours spécifiques du registre du cancer et devront être établis par le registre lui-même. Il faut bien penser à développer un système de codage hiérarchique qui facilitera l'extraction des informations à partir de la base de données du registre (par exemple les listes des cas de cancer) et permettra de retrouver les archives des listes de cas.

Le système de codage pourrait ainsi comporter plusieurs niveaux, comme par exemple :

1. Type de source (hôpital, laboratoire de diagnostic, certificat de décès)
 - 1.1. Liste des hôpitaux – secteur public
 - 1.2. Liste des hôpitaux – secteur privé
 - 1.3. Hospices
 - 1.1.1. Services cliniques (médecine, chirurgie, radiothérapie, etc.).

Il est important, quand on établit un système de codage, de prévoir la possibilité d'une extension future, pour pouvoir inclure de nouvelles sources de données, tout en respectant la structure du système de codage.

Comme cela a été noté plus haut (Tableau 4.2), le registre va inclure le numéro d'enregistrement du cas, mais sauf si l'hôpital ou le service de laboratoire auquel ce numéro se réfère est évident, il sera très difficile de retrouver le dossier si l'on en a besoin pour extraire des informations supplémentaires, pour corriger des

erreurs dans la base de données du registre, ou à des fins de recherche.

5. Informations sur la population à risque

Comme cela a été décrit au Chapitre 3, le registre doit maintenir un fichier de population qui, pour chaque année

calendaire, comprend les estimations pour chaque combinaison de :

- groupe ethnique (si pertinent)
- sexe
- âge (groupes d'âge standard de cinq ans, en séparant, Si possible, les nourrissons [âgés de 0 à 12 mois] des jeunes enfants [âgés de 1 à 4 ans]), et en incluant le nombre de personnes d'âge inconnu.

Points clés

- L'utilisation de multiples sources d'information sur les cas de cancer survenant dans la population cible représente un élément essentiel des registres de population. Les procédures utilisées par les registres permettent d'identifier le même cas de cancer à partir de différentes sources (tout en évitant les doublons). Les sources d'information peuvent être classées en trois grandes catégories : hôpitaux, laboratoires et certificats de décès.
- La plupart des registres associent méthodes actives et méthodes passives pour la recherche des cas.
- Le développement de systèmes informatisés d'information médicale offre la possibilité d'utiliser les bases de données électroniques pour la recherche des cas.
- Les registres du cancer ont pour objectif d'enregistrer, pour chaque cas de cancer, les données correspondant à un certain nombre de variables. On a toujours tendance, lors de la planification du registre, à vouloir inclure trop de variables.
- Il existe 17 à 20 variables essentielles à collecter pour chaque cas enregistré. Il convient de ne rajouter qu'un minimum de variables « optionnelles ». Plusieurs des variables figurant dans la liste doivent être codées pour en faciliter l'analyse. Pour un certain nombre de variables, il existe des systèmes de codage internationaux standardisés, et les registres du cancer doivent les utiliser pour que les résultats des différents registres soient comparables.
- Les systèmes de codage les plus importants concernent le codage de la tumeur (site anatomique, histologie, comportement, base de diagnostic) selon la classification CIM-O et le codage de leur extension anatomique (stade) selon le système TNM.