

Mise au point

## La recherche sur les bases de données en réanimation ☆

### Research on clinical databases in critical care

J.-F. Timsit<sup>a,b,\*</sup>, D. Nakache<sup>c</sup>, C. Alberti<sup>d</sup>, E. Metais<sup>c</sup>

<sup>a</sup> Groupe d'épidémiologie, unité Inserm U578, Grenoble, France

<sup>b</sup> Service de réanimation médicale, CHU A.-Michallon, 38043 Grenoble, cedex France

<sup>c</sup> CNAM, Laboratoire CEDRIC du conservatoire national des arts et métiers, département d'informatique, 292, rue Saint-Martin, 75141 Paris cedex 03, France

<sup>d</sup> Unité d'épidémiologie clinique, hôpital Robert-Debré, 48, boulevard Serrurier, 75019 Paris, France

Disponible sur internet le 11 juillet 2006

#### Résumé

L'avènement de l'informatique dans notre système de santé rend un grand nombre de données médicales disponible pour la recherche. À condition de bien contrôler le mode de saisie, les définitions des variables, et d'effectuer un contrôle de qualité stricte des données il est possible de créer des bases de haute qualité compatible avec une recherche clinique de haute qualité. La recherche sur base de données peut dans ce cas devenir un complément indispensable aux études cliniques randomisées contrôlées. Ces résultats ne sont possibles que sous l'impulsion d'un comité de pilotage multidisciplinaire (réanimateurs, infirmières, biostatisticiens et informaticiens) qui maintient la précision et l'exactitude des données saisies, la motivation des investigateurs, et définit l'orientation de la base et les projets de recherche.

© 2006 Publié par Elsevier SAS pour la Société de réanimation de langue française.

#### Abstract

A huge amount of medical data is now available for clinical research because of the computerization of the healthcare system. If data records are complete and accurate and use standard definitions it becomes possible to develop high quality medical databases and subsequent high quality clinical research. Research on high quality clinical databases should become a useful complement to the randomised controlled trials. However, an important implication of a multidisciplinary (Intensive Care Unit physicians and nurses, biostatisticians) steering committee is needed to maintain the exactness and the completeness of recorded data, to reinforce the motivation of the investigators and to define orientation of the database and future researches.

© 2006 Publié par Elsevier SAS pour la Société de réanimation de langue française.

*Mots clés* : Base de données ; Recherche ; Modélisation ; Qualité ; Informatisation ; Réanimation

*Keywords*: Databases; Clinical research; Modelling; Management audit; Computerization; Critical care

#### 1. Introduction

Ces 20 dernières années, les données médicales sont progressivement informatisées. La plupart des données hospitalières concernant le dossier médical, les prescriptions, l'utilisation des médicaments, des spécialités coûteuses, la réalisation d'ac-

tes de diagnostic, les coûts sont saisis sur un support informatique.

L'informatisation de l'information a abouti à la création progressive d'un grand nombre de bases de données dites opérationnelles dans toutes les activités quotidiennes de l'hôpital (Fig. 1). Les réanimateurs ont eux-mêmes développé des bases de données clinicobiologiques [1–9] leur permettant de connaître la population des patients qu'ils accueillent, de sélectionner un certain nombre de paramètres pouvant constituer le tableau de bord de l'unité (performances, infections nosocomiales, événements iatrogènes) afin de répondre aux besoins

☆ Mise au point réalisée à l'occasion de la journée FMC de La SRLF du 7 avril 2005 consacrée à la recherche clinique en réanimation.

\* Auteur correspondant.

Adresse e-mail : [jf.timsit@outcomerea.org](mailto:jf.timsit@outcomerea.org) (J.-F. Timsit).

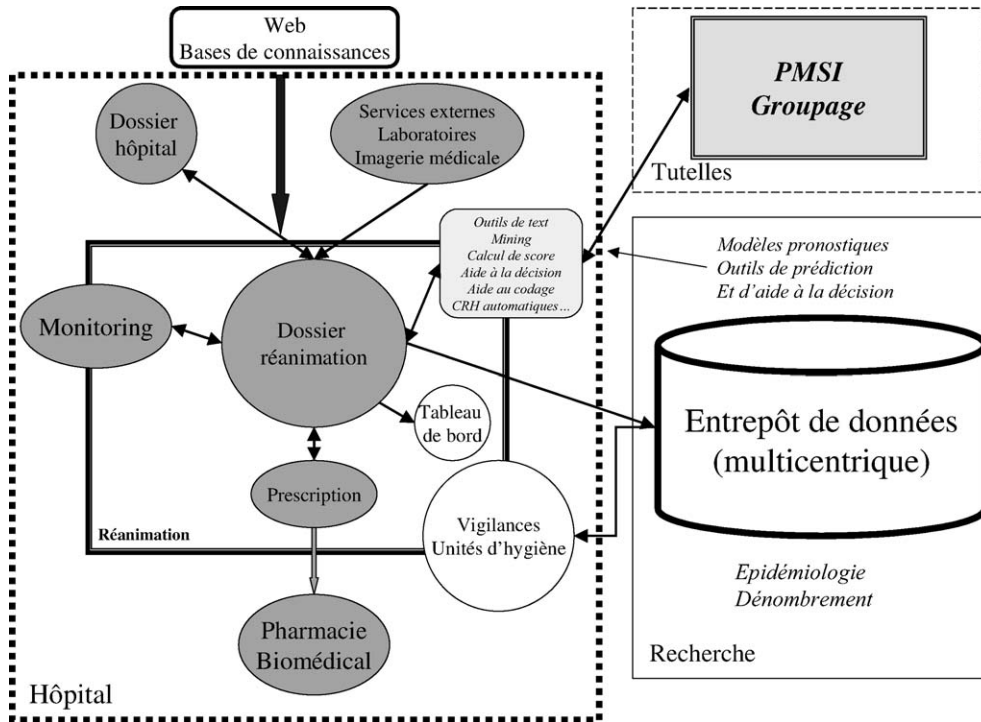


Fig. 1. Articulation des différents systèmes de données disponibles en réanimation.

d'évaluation de l'activité médicale dans les services de réanimation par les administratifs, par les consommateurs ou par les cliniciens, qui ont augmenté de manière considérable.

Il est devenu concevable d'utiliser ces données, dont l'accès est bien plus aisé que ne l'étaient les traditionnelles archives, à des fins de recherche.

## 2. Utilisation des bases de données

### 2.1. De l'évaluation des pratiques à la recherche en épidémiologie

À condition qu'une base de données soit de haute qualité, elle peut être utilisée dans de nombreux domaines [10] :

- en pratique clinique, elle peut fournir, rapidement et sans nécessité de moyens supplémentaires, une estimation des fréquences de différents événements dans différentes populations ;
- elle peut aider au processus décisionnel des praticiens, des patients et/ou de leurs familles ;
- au niveau de l'administration hospitalière, elle permet dans certains cas, une adaptation des besoins aux ressources [11–15] ;
- au-delà de ces missions de routine, ces bases de données de haute qualité peuvent être utilisées pour effectuer des analyses épidémiologiques de sous-groupes. Ces analyses peuvent aboutir à la découverte de facteurs pronostiques ou de facteurs prédictifs d'événements et par là même entraîner la mise en place d'études interventionnelles ou la modification de certains comportements.

Une base de données de haute qualité peut aussi permettre d'auditer les performances des réanimations. À partir du moment où les données sont précises et bien définies, certaines comparaisons d'un centre avec d'autres centres peuvent avoir un sens.

Ces audits peuvent permettre des comparaisons à plus grandes échelles, comme des audits comparatifs entre hôpitaux, groupements d'hôpitaux, régions, etc.

La recherche de facteurs prédisposant à une maladie mais liés au malade lui-même (origine endogène), peut permettre de créer des indicateurs de risque composite d'acquisition d'événements mieux à même d'effectuer des comparaisons entre différentes structures.

Par exemple, le clinicien dans sa routine quotidienne, le manager ou les tutelles peuvent vouloir connaître le risque de développement de certains événements indésirables (par exemple, les infections nosocomiales) sur une certaine sous-population et comparer pour un niveau de risque particulier, le risque observé–risque prédit ou le risque observé dans une unité A par rapport au risque observé dans une unité B chez les patients ayant le même niveau de risque. La mise au point d'indicateurs composites de risques permettant de comparer, toute chose égale par ailleurs, l'unité A avec l'unité B pour la survenue d'événements (décès, acquisition d'infection nosocomiale, acquisition d'événements iatrogènes, satisfactions, etc.) est un des objectifs majeurs de la recherche sur bases de données.

## 3. Recherche sur bases de données et causalité

Les bases de données de haute qualité permettent d'évaluer la prévalence et l'incidence des événements ou des facteurs de

risque par l'intermédiaire d'études de cohorte, d'études cas-témoin ou de séries chronologiques. Le cahier informatique de la base de données peut aussi être utilisé pour réaliser des essais non randomisés.

Ces études ne démontrent pas la causalité entre un facteur de risque et un événement, mais elles sont un complément indispensable et un déclencheur d'essai randomisé contrôlé. En effet, la recherche sur bases de données n'est pas du tout antinomique, bien au contraire, à la réalisation d'essais randomisés [16,17]. La réalisation d'études sur base de données peut stimuler la réalisation d'études randomisées par une meilleure prise de conscience des niveaux d'incertitude de la part des professionnels de santé. Elle facilite la mise en place de l'étude car le cahier de saisie est déjà en place et déjà informatisé. Elle facilite l'inclusion de centres investigateurs, nombreux, et plus proches de la routine quotidienne. Ainsi, la recherche clinique ne devient plus une prérogative systématique de certains centres particuliers hyperspécialisés dotés de moyens spécifiques.

La recherche sur bases de données permet aussi la généralisation des résultats des études randomisées. En effet, on reproche fréquemment aux essais randomisés d'inclure un certain nombre de patients non représentatifs de la population de patients soignés de manière routinière en réanimation [17]. L'existence d'une base de données de haute qualité peut permettre d'effectuer des études épidémiologiques sur une population représentative de la population globale avec un fort taux de participation et un panel large d'hôpitaux, et peut conforter les résultats d'études randomisées et éventuellement encourager l'application de ces résultats dans la pratique.

#### 4. Qui dit base de données de haute qualité, dit qualité des données

Le problème fondamental de la recherche sur base de données est bien entendu la connaissance de la qualité des données saisies. De nombreuses équipes s'efforcent donc de publier des standards de qualité [18–22] et de favoriser la publication sur le Web de la qualité des bases de données existantes [23].

##### 4.1. Quels malades ?

La saisie de l'ensemble des malades admis [20] est recommandée. Si cela est impossible, un échantillon représentatif de la population (constitué par une méthode de tirage au sort) est acceptable. Il faut que la base de données contienne les données essentielles qui permettent de savoir quels types de malades sont hospitalisés « case-mix » et le devenir des patients quels que soient les critères de jugement ou « outcome » choisis [10].

##### 4.2. Quelles variables recueillir ?

Afin d'assurer une bonne qualité des données, avant la mise en route de toute base de données il convient de bien déterminer les variables à recueillir. Cette réflexion préalable doit être approfondie. Il convient d'utiliser des variables objectives et

reproductibles. Il faut limiter leur nombre et si possible éviter les redondances. Dans tous les cas, la standardisation des mesures est un maître mot : il faut des définitions précises et reproductibles de l'ensemble des éléments. Les données recueillies proviennent schématiquement de deux types de source : elles sont soit automatiquement mesurées et transférées à partir des moniteurs et des serveurs de résultats, soit saisies par le personnel médical et paramédical.

##### 4.2.1. La saisie automatique des données en provenance des laboratoires

La saisie automatique des données en provenance des laboratoires par exemple expose à moins d'erreurs de saisie que les retranscriptions manuelles [2,24]. Lorsque certaines données sont récupérées directement à partir de systèmes de surveillance (scope par exemple), il convient de définir le rythme d'enregistrement des mesures (ou granularité) et de déterminer au préalable comment seront validées les données avant intégration à la base de données. Il faut en effet garder à l'esprit que les données recueillies directement par les moniteurs ne sont pas le reflet fidèle de la réalité des données. Par exemple, Tsien et al. [25] ont montré que 92 % des alarmes étaient en fait des faux positifs. Ce taux de faux positif était en plus très variable en fonction des moniteurs. Bien que certains logiciels puissent permettre d'affiner la reconnaissance du signal et la détection plus précise des vraies alarmes, le transfert direct des données des moniteurs vers la base de données sans contrôle et validation est à éviter.

##### 4.2.2. Les données saisies par l'homme

Les données saisies par l'homme représentent une grande part de l'information. Elles concernent tout particulièrement les données complexes comme la saisie des diagnostics médicaux, ou l'interprétation des examens complémentaires. La saisie humaine est aussi très importante en tant que filtre de validation des données fournies automatiquement par les moniteurs.

##### 4.2.3. les bases de données à partir d'un support informatique

À l'heure actuelle, les bases de données doivent être constituées à partir d'un support informatique. Ce support en lui-même améliore la qualité des données recueillies [2]. Il faut donc créer un outil client qui permet d'optimiser l'ergonomie de la saisie, de créer des alarmes de saisie, des contrôles multiples et des menus d'aide interactifs. Par l'intermédiaire d'un accès Internet ou par l'intermédiaire d'un numéro d'appel spécial téléphonique, l'utilisateur doit pouvoir être mis en relation avec le centre de pilotage de la base de données. De même, le centre de pilotage doit pouvoir à tout moment faire parvenir aux utilisateurs de la base de données des détails concernant les variables saisies, la qualité des données, et la réponse à un forum de questions.

Lors de l'initiation de toute base de données, des séances d'entraînement doivent être prévues [26]. Le centre de pilotage doit pouvoir contrôler au mieux la façon dont sont saisies les

données. Ainsi, il est absolument fondamental que l'information soit saisie dès qu'elle est disponible.

#### 4.3. Granularité de la base de données

Les possibilités de la base de données dépendent aussi du niveau de précision disponible. Par exemple, la plupart des bases de données ne stockent l'information agrégée qu'une fois par séjour [1,3,6,9,27], d'autres ont choisi d'intégrer des données pour chaque jour d'hospitalisation [28]. On peut tout à fait envisager de saisir toute l'information (examens biologiques, constantes vitales, événements indésirables etc.) qui se produit à un instant donné. La création de ce type de base de données est d'ailleurs déjà tout à fait possible avec des logiciels commercialisés.

La saisie d'autant de données pose des problèmes majeurs : Comment décider qu'une donnée est correcte et non liée à une erreur de mesure ? que doit-on archiver ? toutes les données (comment gérer la variabilité intra-individuelle ?), une seule (laquelle ? la première, la moyenne, l'aire sous la courbe...) [24] ? Dans tous les cas, cette décision doit être prise avant de débiter la saisie et les hypothèses clairement exposées.

#### 4.4. Qui saisit ou la motivation des investigateurs

##### 4.4.1. La motivation des investigateurs

La motivation des investigateurs est une condition sine qua non pour obtenir une base de données de haute qualité et permettre la recherche [24]. La création de forum d'utilisateurs, la rétro-information concernant les données saisies, les audits, les utilisations potentielles à des fins de routine ou de recherche doivent être organisés.

##### 4.4.2. La qualité de l'interface utilisateur

La qualité de l'interface utilisateur peut être utilisée en tant qu'outil de motivation. Ainsi, il est possible de créer des interfaces permettant de saisir non seulement les données utiles pour la base de données, mais aussi des données utiles pour l'hôpital ou le service (moteur de recherche : intégrés des codes de la CIM-10 ou des actes CCAM, tableau de bord informatique concernant l'activité du service ou de l'hôpital, génération informatisée des comptes-rendus d'hospitalisation, génération et export automatisé des actes classants, intégration de pages données réservées à l'utilisation et non exportées vers la base de données).

##### 4.4.3. Les utilisateurs de la base de données

Enfin, les utilisateurs de la base de données doivent pouvoir l'utiliser à des fins de recherche. Un accord préalable doit être passé entre les utilisateurs et le comité de pilotage sur les conditions d'utilisation des données et les conditions de publication. À cette condition, les bases de données permettent de générer de nombreuses études épidémiologiques [29–35] donnant lieu à une publication dans des revues parfois prestigieuses.

#### 4.5. Archivage des données recueillies

Plusieurs étapes de contrôle sont fondamentales avant archivage des données. En plus des alarmes de saisie initiale, les données archivées doivent être soumises à des programmes de contrôle des incohérences potentielles [36] et de corrections automatiques ou manuelles. L'ensemble des erreurs doit être soumis en temps réel aux utilisateurs à des fins de correction avant archivage. Cette rétro-information au centre peut en elle-même améliorer la qualité de la base de données [37]. Les audits périodiques sont indispensables. Ces audits sont coûteux, mais doivent donner lieu à des analyses poussées et à un rapport qui doit être pris en compte [2,4,8,38].

### 5. Vie d'une base de données

Une base de données de haute qualité utilisable en recherche est vivante. Dès que l'outil de saisie est abandonné, les données vieillissent, les définitions sont oubliées, et après quelques années toute utilisation à des fins de recherche devient impossible.

Si la base de données reste vivante, il faudra s'attacher à la structure des données archivées, en particulier à l'historisation de ces données et à l'historisation des variables saisies.

Un comité multidisciplinaire de pilotage doit être créé. Ce comité de pilotage est fondamental pour assurer l'aspect politique de la base de données (gestion financière, sécurité...) et l'aspect technique [10] :

- informatique (contrôle avant archivage des données, sauvegarde des données des différents centres, sécurité des données, cryptage) ;
- épidémiologique et biostatistique (méthodologie des études épidémiologiques et des essais, analyse statistique, écriture des articles).

Ce comité de pilotage déterminera les orientations de la base de données et doit avoir pour obsession permanente le maintien de la qualité des données saisies.

Il doit assurer la pérennité d'une plate-forme de communication :

- la création d'un site Internet reprenant les différents référentiels permettant un forum de questions, l'archivage, le cryptage des données, la mise en ligne des rapports d'activité ;
- des réunions d'utilisateurs périodiques ;
- des réponses aux demandes des différents centres ;
- un bulletin périodique d'information sur l'état de la base de données.

### 6. Base de données : peut-on tout faire ?

#### 6.1. Les objectifs de la recherche sur base de données

La recherche qui va être réalisée sur les bases de données est extrêmement dépendante des objectifs initiaux de cette

base. Il faut en effet bien distinguer les bases de données orientées exclusivement vers la recherche et développant certains bénéfices secondaires pour les utilisateurs et la recherche sur base de données dont le but initial n'est pas la recherche mais une utilisation à des fins administratives, comptables, médicales ou commerciales [39].

Le regroupement de données provenant de plusieurs bases de données hospitalières dans un dossier clinique informatique peut parfois entraîner des erreurs : [24]. L'objectif initial de la base de données va en effet conditionner le type de variables saisies et la façon dont elles sont saisies. Par exemple, la façon dont sont saisis les codes CIM-10 afin de valoriser les séjours hospitaliers (PMSI) est très différente de la façon dont les médecins réanimateurs qualifieraient le séjour hospitalier de leurs patients. Ou encore, l'utilisation des glycopeptides pour traiter des infections à staphylocoque résistant à la méthicilline utilise des données très différentes pour les pharmaciens (nombre de grammes de vancomycine par unités de temps utilisées dans les services ou dose définie journalière) et pour les réanimateurs (dose de vancomycine en milligramme par kilogramme, vancocinémie, au pic, à la vallée...) [40].

Avant d'entreprendre une recherche sur une base de données, il faut bien se poser la question « Pour quel objectif principal les variables qui m'intéressent sont-elles saisies ? » [10]. Les bases de données constituées peuvent en effet privilégier les coûts, les médicaments, le matériel, les infections nosocomiales, les événements indésirables, la performance hospitalière, elles peuvent aussi privilégier le malade, pathologie, diagnostic, score de gravité, dysfonction d'organes généraux ou locaux.

## 6.2. Limites de la recherche sur base de données

- Bien entendu, la recherche sur base de données a des limites. Certaines données ou événements rares peuvent être plus facilement oubliés ou mal saisis [41] ;
- au contraire, certains investigateurs peuvent être particulièrement sensibilisés à la saisie de certains paramètres, au détriment d'autres. Par exemple, la saisie des actes marqueurs de réanimation sera probablement plus précise et plus reproductible que la saisie des autres actes classants ;
- un exemple marquant de l'absence de reproductibilité des données concerne les diagnostics. Malgré l'utilisation préférentielle d'un thésaurus CIM-10 adapté à la réanimation, des recommandations de codage effectuées par les sociétés savantes, la reproductibilité des diagnostics codés pour un même séjour est très mauvaise. Ainsi, nous avons fait recoder en aveugle par deux réanimateurs seniors 100 séjours de réanimation. Seuls 18 % des codes diagnostiques générés étaient retrouvés à la fois par le codage initial et par les deux recodeurs. La reproductibilité était très variable selon les codes, excellente pour certains (ex : coma 100 % de consensus) médiocre pour d'autres (ex : insuffisance respiratoire aiguë 19 %, pneumopathie bactérienne 27 %) [42]. Des résultats tout aussi médiocres ont été rapportés par des équipes nord-américaines [24].

Une autre difficulté de la recherche sur base de données, si grande soit-elle, est la représentativité de cette base de données. En effet, même dans les grandes bases de données, l'échantillon n'est pas homogène et pas forcément représentatif de la population. Par exemple, dans la base de données ICNARC développée au Royaume-Uni [1], le nombre d'admissions pour dissection ou anévrisme de l'aorte ou de l'artère iliaque représente 7406 séjours, soit près de 6 % des admissions en réanimation. Il est peu vraisemblable que la population des malades de réanimation soit constituée d'un aussi grand pourcentage de tels malades. Si l'on compare les principales bases de données européennes ou américaines, le pourcentage de chirurgie programmée varie de moins de 10 % [4] à plus de 65 % [9] ; la mortalité hospitalière varie de moins de 18 % à plus de 30 % [24].

## 7. La base de données de haute qualité est efficace mais coûteuse

La mise en place et le développement d'une base de données permettant la recherche ont un coût élevé mais sont coût-efficaces. En effet, les systèmes de saisie sont déjà en place et les coûts sont répartis sur plusieurs études et partagés avec d'autres implications (pratique clinique, dossier médical, management ou audit).

La recherche de moyens financiers pérennes est un objectif quotidien du comité de pilotage. Une partie de la base peut être financée par les structures hospitalières ou administratives. La base peut également être financée par des fonds de recherche.

Cependant, la recherche médicale ne favorise pas le financement des bases de données. Très peu d'appels d'offres publics sont destinés à financer des bases de données. Certains partenaires industriels peuvent être intéressés, cependant les partenariats doivent définir très précisément la propriété des données, l'utilisation potentielle de ces données à des fins non plus de recherche, mais industrielles et la confidentialité de ces données.

Il faut bien garder à l'esprit que le financeur de la base de données peut modifier l'objectif de la base de données [10]. Le comité de pilotage devra choisir entre un financeur unique pérenne pouvant modifier l'esprit de la base de données de recherche, et des financeurs multiples permettant d'éviter ce risque mais dont la volatilité accrue met en danger l'avenir même de l'instrument.

## 8. Base de données de recherche : faut-il augmenter sa taille en augmentant l'assiette de saisie ?

L'augmentation de la taille de la base de données permet d'augmenter probablement la représentativité du *case-mix* et l'exhaustivité de l'échantillon dans la population [20]. Cependant, cette augmentation entraîne l'augmentation du nombre d'investigateurs, donc de codeurs, une augmentation du nombre de personnes non directement impliquées dans la base de données participant à la saisie, et donc une augmentation de la

variabilité des données saisies. Bien entendu, l'augmentation de taille entraîne aussi une augmentation des coûts.

Il s'agira de trouver le bon équilibre entre un nombre suffisant de centre pour assurer la représentativité tout en restant raisonnable pour en maîtriser la complexité.

La constitution d'entrepôts de données (*data warehouses*) est une réponse au problème de l'intégration d'une grande quantité de données variées, relatives à un certain domaine d'application, et stockées physiquement dans différentes sources de données. L'entrepôt de données regroupe sous une forme exploitable les informations extraites de ces sources et qui sont potentiellement pertinentes pour les décideurs.

Un entrepôt de données est défini comme une collection de données intégrées, orientées sujettes, historisées et disponibles pour l'interrogation, l'analyse et la prise de décision. Initialement les entrepôts de données étaient destinés à la prise de décision dans le milieu commercial mais les principes de l'aide à la décision peuvent également s'appliquer avec succès au domaine médical.

L'entrepôt de données est au cœur du système décisionnel : il va permettre de disposer d'un stock de données médicales nettoyées, intégrées et historisées. On pourra poser à ce gisement de données des requêtes hautement complexes (par exemple rechercher la molécule ayant eu le plus grand impact positif sur un symptôme, avec un impact négatif sur un autre symptôme ne dépassant pas telle durée) ou partir à la recherche de corrélations inconnues entre les données (par exemple des relations de cause à effet entre deux pathologies, même si les événements corrélés se sont produits avec un intervalle de temps). Citons parmi les applications possibles l'analyse des risques iatrogènes ou la formulation de requêtes épidémiologiques.

Les données provenant d'hôpitaux différents doivent subir une étape d'intégration afin de résoudre les problèmes d'hétérogénéité de vocabulaire (ex : différentes définitions de l'infection nosocomiale selon les pays) ou de pratiques (ex : analyse bactériologique avant ou après antibiothérapie).

Une architecture fondée sur un entrepôt de données garantit la confidentialité des données relatives aux patients et aux établissements. Cette caractéristique est gérée au travers de la dualité entre « données locales » et « données publiques ». Chaque hôpital reste propriétaire de ses données dans sa propre base. Une partie des données sont dupliquées dans l'entrepôt commun après anonymisation.

## 9. Conclusion

Les bases de données existantes et en cours de constitution sont des outils qui ont d'énormes potentiels en termes de recherche. La recherche sur base de données est très certainement un complément, voire même un soutien au développement d'études randomisées contrôlées. Bien entendu, la recherche sur base de données est conditionnée par la qualité des données saisies, et certaines études réalisées sur d'énormes quantités de données sont finalement vaines parce que les définitions utilisées ne sont pas précises et reproductibles.

## Références

- [1] Harrison DA, Brady AR, Rowan K. Case mix, outcome and length of stay for admissions to adult, general critical care units in England, Wales and Northern Ireland: the Intensive Care National Audit and Research Centre Case Mix Programme Database. *Crit Care* 2004;8:99–111.
- [2] Cook SF, Visscher WA, Hobbs CL, Williams RL. Project IMPACT: results from a pilot validity study of a new observational database. *Crit Care Med* 2002;30:2765–70.
- [3] Valentin A, Jordan B, Lang T, Hiesmayr M, Metnitz PG. Gender-related differences in intensive care: a multiple-center cohort study of therapeutic interventions and outcome in critically ill patients. *Crit Care Med* 2003; 31:1901–7.
- [4] Aegerter P, Auvert B, Buonamico G, Sznajder M, Beauchet A, Guidet B, et al. Mise en oeuvre et évaluation d'une base de données commune aux services de réanimation d'Île-de-France. *Rev Epidemiol Sante Publique* 1998;46:226–37.
- [5] Timsit JF, Fosse JP, Troche G, De Lassence A, Alberti C, Garrouste-Orgeas M, et al. Accuracy of a composite score using daily SAPS II and LOD scores for predicting hospital mortality in ICU patients hospitalized for more than 72 hours. *Intensive Care Med* 2001;27:1012–21.
- [6] Moreno R, Apolone G, Miranda DR. Evaluation of the uniformity of fit of general outcome prediction models. *Intensive Care Med* 1998;24:40–7.
- [7] Sirio CA, Tajimi K, Taenaka N, Ujike Y, Okamoto K, Katsuya H. A cross-cultural comparison of critical care delivery: Japan and the United States. *Chest* 2002;121:539–48.
- [8] Arts D, de Keizer N, Scheffer GJ, de Jonge E. Quality of data collected for severity of illness scores in the Dutch National Intensive Care Evaluation (NICE) registry. *Intensive Care Med* 2002;28:656–9.
- [9] de Jonge E, Bosman RJ, van der Voort PH, Korsten HH, Scheffer GJ, de Keizer NF. Intensivecaregeneeskunde in Nederland, 1997–2001. I. Patientenpopulatie en resultaten van behandeling. *Ned Tijdschr Geneesk* 2003;147:1013–7.
- [10] Black N. High-quality clinical databases: breaking down barriers. *Lancet* 1999;353:1205–6.
- [11] Rowan KM, Kerr JH, Major E, McPherson K, Short A, Vessey MP. Intensive Care Society's Apache II study in Britain and Ireland—II: outcome comparisons of intensive care units after adjustment for case mix by the American Apache II method. *BMJ* 1993;307:977–81.
- [12] Carmel S, Rowan K. Variation in intensive care unit outcomes: a search for the evidence on organizational factors. *Curr Opin Crit Care* 2001;7: 284–96.
- [13] Goldfrad C, Rowan K. Consequences of discharges from intensive care at night. *Lancet* 2000;355:1138–42.
- [14] Angus DC, Black N. Improving care of the critically ill: institutional and health-care system approaches. *Lancet* 2004;363:1314–20.
- [15] van Walraven C, Demers SV. Coding diagnoses and procedures using a high-quality clinical database instead of a medical record review. *J Eval Clin Pract* 2001;7:289–97.
- [16] Padkin A, Rowan K, Black N. Using high quality clinical databases to complement the results of randomised controlled trials: the case of recombinant human activated protein C. *BMJ* 2001;323:923–6.
- [17] McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C. Methods in health services research. Interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ* 1999;319:312–5.
- [18] Mothral B, Brooks J, Clark MA, Crown WH, Davey P, Hutchins D, et al. A checklist for retrospective database studies—report of the ISPOR Task force on retrospective databases. *Value Health* 2003;6:90–7.
- [19] Whiting P, Rutjes AW, Dinnes J, Reitsma J, Bossuyt PM, Kleijnen J. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004;8(3):1–234.
- [20] Proctor SJ, Taylor PR. A practical guide to continuous population-based data collection (PACE): a process facilitating uniformity of care and research into practice. *QJM* 2000;93:67–73.

- [21] Alberti C, Timsit JF, Porcher R, Chevret S. Contrôle de qualité des données. In: Bion J, editor. *Management en réanimation: évaluation, organisation et éthique*. Paris: Elsevier SAS; 2000. p. 112–30.
- [22] Arts DG, De Keizer NF, Scheffer GJ. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. *J Am Med Inform Assoc* 2002;9:600–11.
- [23] Black N, Payne M. Directory of clinical databases: improving and promoting their use. *Qual Saf Health Care* 2003;12:348–52.
- [24] Ward NS. The accuracy of clinical information systems. *J Crit Care* 2004;19:221–5.
- [25] Tsien CL, Fackler JC. Poor prognosis for existing monitors in the intensive care unit. *Crit Care Med* 1997;25:614–9.
- [26] Arts DG, Bosman RJ, de Jonge E, Joore JC, de Keizer NF. Training in data definitions improves quality of intensive care data. *Crit Care* 2003;7:179–84.
- [27] Aegerter P, Boumendil A, Retbi A, Minvielle E, Dervaux B, Guidet B. SAPS II revisited. *Intensive Care Med* 2005;31:416–23.
- [28] Timsit JF, Fosse JP, Troche G, De Lassence A, Garrouste-Orgeas M, et al. Calibration and discrimination by daily logistic organ dysfunction scoring comparatively with daily sequential organ failure assessment scoring for predicting hospital mortality in critically ill patients. *Crit Care Med* 2002;30:2003–13.
- [29] Azoulay E, Adrie C, De Lassence A, Pochard F, Moreau D, Thiery G, et al. Determinants of postintensive care unit mortality: a prospective multicenter study. *Crit Care Med* 2003;31:428–32.
- [30] de Lassence A, Alberti C, Azoulay E, Le Miere E, Cheval C, Vincent F, et al. Impact of unplanned extubation and reintubation after weaning on nosocomial pneumonia risk in the intensive care unit: a prospective multicenter study. *Anesthesiology* 2002;97:148–56.
- [31] Garrouste-Orgeas M, Troche G, Azoulay E, Caubel A, de Lassence A, Cheval C, et al. Body mass index. An additional prognostic factor in ICU patients. *Intensive Care Med* 2004;30:437–43.
- [32] Timsit JF, Fosse JP, Troche G, De Lassence A, Alberti C, Garrouste-Orgeas M, et al. Calibration and discrimination by daily logistic organ dysfunction scoring comparatively with daily sequential organ failure assessment scoring for predicting hospital mortality in critically ill patients. *Crit Care Med* 2002;30:2003–13.
- [33] Annane D, Aegerter P, Jars-Guincestre MC, Guidet B. Current epidemiology of septic shock: the CUB-Rea Network. *Am J Respir Crit Care Med* 2003;168:165–72.
- [34] Guidet B, Aegerter P, Gauzit R, Meshaka P, Dreyfuss D. Incidence and impact of organ dysfunctions associated with sepsis. *Chest* 2005;127:942–51.
- [35] Bornstain C, Azoulay E, De Lassence A, Cohen Y, Costa MA, Mourvillier B, et al. Sedation, sucralfate, and antibiotic use are potential means for protection against early-onset ventilator-associated pneumonia. *Clin Infect Dis* 2004;38:1401–8.
- [36] Svolba G, Bauer P. Statistical quality control in clinical trials. *Control Clin Trials* 1999;20:519–30.
- [37] De Lusignan S, Stephens PN, Adal N, Majeed A. Does feedback improve the quality of computerized medical records in primary care? *J Am Med Inform Assoc* 2002;9:395–401.
- [38] Tafflet M, Alberti C. Audit de la base de données Outcomerea, le rapport. 2004. <http://www.outcomerea.org/ehm/repro-base2003.pdf>.
- [39] Moine P, Hemery F, Bleriot JP, Fulgencio JP, Garrigues B, Gouzes C, et al. Exhaustivité des résumés d'unité médicale adresses aux départements d'information médicale. *Ann Fr Anesth Reanim* 2004;23:15–20.
- [40] Mandy B, Kounty E, Cornette C, Wornoff-Lemsi M-C, Talon D. Comparaison de deux méthodes de mesure de la consommation d'antibiotiques. *Hygiène* 2000;11:465–70.
- [41] Grover FL, Hammermeister KE, Shroyer AL. Quality initiatives and the power of the database: what they are and how they run. *Ann Thorac Surg* 1995;60:1514–21.
- [42] Misset B, Nakache D, Dumont S, De Lassence A, Darmon M, Garrouste-Orgeas M, et al. Reproductibilité des codages diagnostiques en réanimation — projet CIREA 1<sup>re</sup> partie — pour Outcomerea. *Réanimation* 2005;14:S109.