



HAL
open science

Recherche clinique et recueil de données lors de l'investigation : influence d'un personnel dédié

Sabine Ghenim, Thomas Féron, Coralie Barbe, Aurore Wolak-Thierry,
Damien Jolly

► **To cite this version:**

Sabine Ghenim, Thomas Féron, Coralie Barbe, Aurore Wolak-Thierry, Damien Jolly. Recherche clinique et recueil de données lors de l'investigation : influence d'un personnel dédié. *Thérapie*, 2018, 73 (3), pp.267 - 272. 10.1016/j.therap.2017.10.003 . hal-01919177

HAL Id: hal-01919177

<https://hal.science/hal-01919177>

Submitted on 12 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Disponible en ligne sur

ScienceDirect
www.sciencedirect.com

Elsevier Masson France

EM|consulte
www.em-consulte.com



PHARMACOLOGIE CLINIQUE

Recherche clinique et recueil de données lors de l'investigation : influence d'un personnel dédié

Clinical research and data collection during the investigation: Influence of a dedicated staff

Sabine Ghenim^a, Thomas Féron^a, Coralie Barbe^{b,c},
Aurore Wolak-Thierry^{b,*c}, Damien Jolly^{b,c}

^a Centre de recherche et d'investigation clinique, centre hospitalier universitaire de Reims, avenue du Général-Koenig, 51095 Reims, France

^b Unité d'aide méthodologique, pôle recherche et santé publique, centre hospitalier universitaire de Reims, hôpital Robert-Debré, 51095 Reims cedex, France

^c EA 3797, faculté de médecine, université de Reims Champagne-Ardenne, 51095 Reims, France

Reçu le 14 mars 2016 ; accepté le 12 octobre 2017

Disponible sur Internet le 11 novembre 2017

MOTS CLÉS

Données
manquantes ;
ARC

Résumé

Objectif. — En recherche clinique, le recours à du personnel dédié, type attaché de recherche clinique (ARC), constitue une pratique courante visant à faciliter le recueil de données et limiter les données manquantes (DM). Nous avons étudié l'intérêt du recours à un ARC pour le recueil de données.

Méthodes. — Les DM ont été classées selon cinq catégories : cliniques, para-cliniques, traitements, événements indésirables et autres. Les études ont été séparées selon leur design (« sans suivi » pour les études transversales avec une seule visite et « avec suivi » pour les études longitudinales comprenant plusieurs visites) et selon l'organisation du recueil (« sans ARC » si les données étaient recueillies directement par un investigateur et « avec ARC » si les données étaient recueillies par un ARC).

Résultats. — La présence d'un ARC permet de réduire le pourcentage de DM quel que soit leur type (test de Student : $p < 0,0001$) : en moyenne avec un ARC $4,8 \% \pm 8,4 \%$ de DM et sans ARC $22,1 \% \pm 17,0 \%$.

* Auteur correspondant. Centre de recherche et d'investigation clinique, centre hospitalier universitaire de Reims, avenue du Général-Koenig, 51095 Reims, France

Adresse e-mail : awolak-thierry@chu-reims.fr (A. Wolak-Thierry).

Conclusion. – La délégation du recueil des données à un personnel spécialisé permet de diminuer significativement le pourcentage de données manquantes.

© 2017 Société française de pharmacologie et de thérapeutique. Publié par Elsevier Masson SAS. Tous droits réservés.

KEYWORDS

Missing data;
CRA

Summary

Objective. – In clinical research, the recourse to a staff dedicated as CRA with the investigator is common practice to facilitate the work of collection and limiting the missing data (MD). We, therefore, looked for the interest of the recourse of such personnel to data collection.

Methods. – MD were classified according to five categories: clinical, para-clinical, treatment, adverse events (serious) and others. Studies were separated in two designs, one-off studies during a single visit (so-called “no follow”) and studies on the duration and including several visits (say “with follow”). Similarly, studies were differentiated according to their type of collection “Without ARC” if the data were collected by an investigator, and studies “With ARC”.

Results. – The presence of a CRA can reduce the number of MD whatever their type (Student test: $P < 0.0001$): With CRA mean of MD is $4.8\% \pm 8.4\%$ and Without CRA mean of MD is $22.1\% \pm 17.0\%$.

Conclusion. – The delegation of data collection to a staff dedicated reduces significantly the percentage of missing data.

© 2017 Société française de pharmacologie et de thérapeutique. Published by Elsevier Masson SAS. All rights reserved.

Abréviations

ARC	attaché de recherche clinique
CRF	case report form
DM	dispositifs médicaux
ICH	international conference on harmonisation of technical requirements for registration of pharmaceuticals for human use
TEC	technicien d'études cliniques
SC	study coordinators

Introduction

Les études cliniques sont organisées pour faire émerger de nouvelles thérapeutiques, améliorer les techniques diagnostiques, l'usage des traitements existants, prendre des décisions médico-économiques ou enrichir les connaissances scientifiques [1]. Les conclusions des études cliniques peuvent influencer la pratique médicale, voire motiver des décisions des autorités de santé publique. La rigueur dans leur conduite est donc un impératif à respecter. Celle-ci est très encadrée, fondée sur des obligations légales, sur les bonnes pratiques cliniques, des directives internationales; *international conference on harmonisation of technical requirements for registration of pharmaceuticals for human use* (ICH), des lignes directives émises par certaines organisations internationales [2] basées sur les grands

principes fondateurs de la déclaration d'Helsinki et de ses différentes mises à jour.

De plus, il est essentiel de disposer d'un recueil de données de qualité afin de pouvoir prêter du crédit aux conclusions émises. Toutefois, même si la fiabilité des données cliniques est garantie par le monitoring, l'exhaustivité reste exceptionnelle. Usuellement, une partie des données sera manquante malgré la bonne volonté et le soin apporté par tous les intervenants [3,4]. Des méthodes statistiques pour imputer des données manquantes « a posteriori » (méthodes d'imputation de données) existent. Cependant, aucune n'est parfaite et ces méthodes statistiques ne sont applicables que pour des taux faibles de données manquantes. La mise en place de procédures et des formations détaillées, permettent de minimiser « a priori » la présence des données manquantes [1,2].

L'augmentation de la quantité de données requises dans les études cliniques au fil des années a entraîné l'émergence d'une catégorie de personnel spécialisé, dédié à l'assistance des investigateurs et à la saisie des données. Ce sont les techniciens d'études cliniques (TEC), ou attachés de recherche clinique hospitaliers (ARC) désignés en anglais comme *study coordinators* (SC) qui recueillent des données cliniques par délégation et sous la responsabilité de l'investigateur. En disposant d'un temps dédié et d'une spécialisation dans le recueil de données, ils facilitent la conduite des études cliniques. Toutefois, cette aide au recueil de données n'a pas fait l'objet d'évaluation. À notre connaissance, la différence concernant le nombre de données manquantes entre un recueil de données effectué par

un investigateur disposant de connaissances de la pathologie et du patient, mais de peu de temps dédié, et un recueil de données réalisé par un ARC/TEC/SC avec des connaissances théoriques inférieures, mais un temps dédié significatif n'a pas été étudiée à ce jour.

L'objectif principal de cette étude était donc de comparer le pourcentage de données manquantes en fonction de la présence ou non d'un ARC/TEC/SC, en prenant en compte la catégorie des données et le type d'étude.

Méthode

Recrutement

Cette étude a été menée parmi les études en cours au CHU de Reims entre le 01/01/2010 et le 31/01/2015. Les études à promotion industrielle n'étaient pas éligibles. Les études cliniques incluses pouvaient avoir des méthodologies différentes (monocentrique ou multicentrique, observationnelle ou interventionnelle, transversale avec une seule visite par patient ou longitudinale avec plusieurs visites par patient, promues ou non par le CHU de Reims) et pouvaient porter sur toutes les spécialités médicales ou chirurgicales. L'unité statistique de cette étude était la visite-patient. Il n'était pas déterminé le nombre minimal de données à recueillir par visite-patient dans le cahier d'observation (ou *case report form* [CRF]) pour que celle-ci soit prise en compte dans cette étude. Par contre, pour être éligible, les études devaient comporter au moins 20 visites-patient. Les annotations, données répétitives d'identification des pages ou du patient, n'étaient pas comptabilisées.

Description des données recueillies

Les données à recueillir dans le CRF ont été classées en cinq catégories :

- les données « cliniques » comprenant les données biométriques (âge, poids, fréquence cardiaque, tension artérielle), l'histoire de la maladie, les antécédents du patient et les cotations et scores selon des échelles spécifiques ou non à la pathologie (questionnaires de qualité de vie);
- les données « para-cliniques » comprenant les données biologiques et les résultats des examens d'imagerie;
- les données « traitement » regroupant les informations relatives au traitement à l'étude (lorsque celui-ci existait) et les données relatives aux traitements concomitants;
- les données « événements indésirables » comprenant notamment le caractère grave ou non, la date de survenue, le critère d'imputabilité;
- les données « Autres » regroupant les quelques items propres à certains CRF, tels que la profession du patient ou sa participation concomitante à un autre protocole d'essai clinique.

Une donnée était considérée comme manquante du fait de son absence lors du rendu final des données au promoteur.

Les études cliniques ont été classées selon deux catégories :

- les études dites « avec suivi », comprenant une visite d'inclusion, un nombre variable de visites de suivi, et une visite de fin d'étude ;
- les études dites « sans suivi » comprenant une unique visite, le patient n'étant vu qu'une seule fois pour le protocole.

Le mode de recueil de données a été classé selon deux types :

- « Sans ARC/TEC/SC » si les données étaient recueillies directement par un investigateur, sans aucune intervention d'un ARC/TEC/SC;
- « Avec ARC/TEC/SC » si les données étaient transcrites dans les CRF par un personnel dédié.

Taille de l'échantillon

Le calcul du nombre de visites-patient nécessaire a été effectué en utilisant la précision souhaitée pour l'estimation du pourcentage de données manquantes. En prenant un pourcentage de données manquantes attendu de 15 % et une précision souhaitée de ± 5 %, le nombre de visites-patient nécessaire était de 200 (calcul grâce à la formule d'intervalle de confiance de 95 %) :

$$1.96 \times \sqrt{(p \cdot q) / n}$$

Au total, 10 études cliniques avec un recueil de données « Sans ARC/TEC/SC » et 121 études cliniques avec un recueil de données « Avec ARC/TEC/SC » étaient éligibles pour cette étude.

L'ensemble des études cliniques avec un recueil de données « Sans ARC/TEC/SC » ont été incluses. Un tirage au sort a été réalisé afin de déterminer les 200 visites-patient analysées pour cette étude.

Un tirage au sort a été réalisé afin de ne retenir que 10 études cliniques parmi les 121 avec un recueil de données « Avec ARC/TEC/SC ». Un second tirage au sort a ensuite été réalisé afin de déterminer les 200 visites-patient (parmi les 10 études cliniques avec un recueil de données « Avec ARC/TEC/SC ») analysées pour cette étude.

Analyse statistique

Pour chaque visite-patient, le nombre de données manquantes a été rapporté au nombre de données totales demandées. Le test de Student a permis d'étudier la différence de données manquantes en fonction de la présence ou non d'un ARC/TEC/SC puis en fonction du type d'étude (avec ou sans suivi). Ensuite une analyse multivariée par régression linéaire généralisée a été réalisée pour chaque type de données manquantes. Les variables proposées au modèle multivariée devaient présenter un $p < 0,2$ en analyse univariée. Le seuil de signification était de 0,05. L'ensemble des analyses a été réalisé avec le logiciel SAS 9,4 (SAS Inc., Cary, North California).

Résultats

Les résultats de l'analyse univariée comparant les données manquantes en fonction de la présence ou non d'un

ARC/TEC/SC sont présentés dans le [Tableau 1](#) et ceux comparant les données manquantes en fonction du type d'étude (avec ou sans suivi) sont reportés dans le [Tableau 2](#).

En analyse univariée, le pourcentage de données manquantes était significativement inférieur dans les études cliniques « avec ARC/TEC/SC » et ce quel que soit le type de variable ($p < 0,03$). Les pourcentages de données manquantes pour les variables de type « cliniques » et de type « para-cliniques » étaient inférieurs pour les études « avec suivi » ($p < 0,0001$ pour les deux).

En analyse multivariée ([Tableau 3](#)), la présence d'un ARC/TEC/SC ($4,8 \pm 8,4$ versus $22,1 \pm 17,0$, $p < 0,0001$) et le type d'étude avec suivi ($11,8 \pm 16,0$ versus $16,7 \pm 15,5$, $p = 0,01$) influençaient significativement le pourcentage de données manquantes. De plus, il a été retrouvé une interaction quantitative entre la présence d'un ARC/TEC/SC et le type d'étude ($p = 0,03$) ([Fig. 1](#)).

Discussion

L'objectif principal de notre étude était de comparer le pourcentage de données manquantes en fonction de la présence ou non d'un ARC/TEC/SC, en prenant en compte la catégorie des données et le type d'étude. Les résultats indiquent que, quel que soit le type de données observées, la délégation du recueil des données à un personnel spécialisé permet de diminuer significativement le pourcentage de données manquantes. Ceci peut s'expliquer par le fait que l'ARC/TEC/SC dispose d'un temps dédié et d'une disponibilité permettant une recherche des données plus approfondie dans les documents sources. Les données manquantes peuvent également être dues à l'absence de ces

données dans les documents source. L'ARC/TEC/SC peut alors effectuer un travail de recherche de la donnée absente des documents source, mais récupérable. Il faut aussi noter que l'ARC/TEC/SC peut être présent lors des visites-patient et veiller à la transcription par l'investigateur des données requises par l'étude dans les documents-sources.

Ensuite nous cherchions à comparer le pourcentage de données manquantes en fonction du type d'étude (avec ou sans suivi). Le pourcentage de données manquantes des études « avec suivi » est significativement plus faible que celui des études « sans suivi ». En effet, la multiplicité des visites permet de revenir sur une donnée manquante à l'occasion d'une visite suivante. La saisie régulière des données dans le CRF permet de mettre en évidence les données manquantes et de les cibler lors des visites ultérieures. Toutefois, si la différence existait pour la variable DM totale, l'analyse par type de données révélait que seules les données cliniques et para-cliniques étaient concernées. Concernant les données relatives aux traitements, l'absence de différence s'explique par la nature des documents-source, presque toujours présents et disponibles en raison des obligations légales, des procédures standardisées et des bonnes pratiques de pharmacie.

Enfin, l'analyse multivariée a montré une interaction quantitative entre la présence d'un ARC/TEC/SC et le type d'étude. En effet, cela signifie qu'en présence d'un ARC, on retrouve autant de DM quel que soit le type de suivi, mais lorsqu'il n'y a pas d'ARC, le pourcentage de DM est beaucoup plus élevé dans les études avec des visites de suivi par rapport aux études qui ne comportent qu'une seule visite.

À notre connaissance, notre étude fait partie des rares publications retrouvées sur le sujet. Un autre de ses

Tableau 1 Pourcentage de données manquantes (DM) en fonction de la présence ou non d'un TEC (test de Student).

Types variables (%)	Sans TEC (n = 200)	Avec TEC (n = 200)	p
DM cliniques	12,8 ± 18,3	3,2 ± 9,8	< 0,0001
DM paracliniques	23,5 ± 23,7	7,8 ± 17,9	< 0,0001
DM traitements	26,9 ± 28,1	5,4 ± 14,9	< 0,0001
DM EI/EIG	22,69 ± 35,91	2,24 ± 8,94	0,0279
DM autres	7,1 ± 17,9	2,5 ± 12,3	0,01
DM totales	22,1 ± 17,0	4,8 ± 8,4	< 0,0001

EI/EIG : évènements indésirables/évènements indésirable graves.

Tableau 2 Pourcentage de données manquantes (DM) en fonction du type d'étude (test de Student).

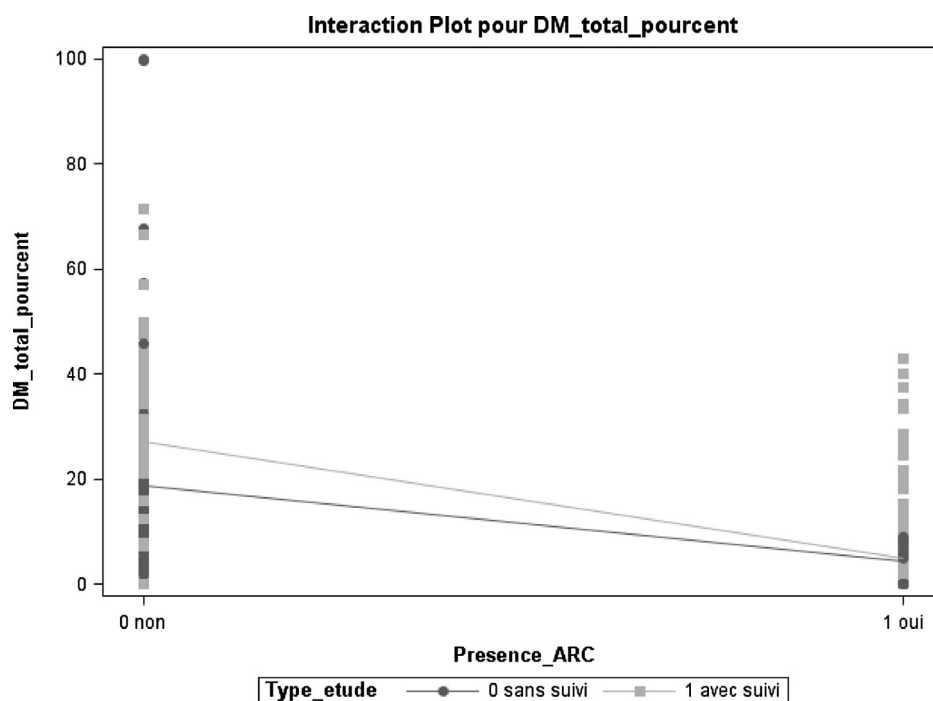
Types variables (%)	Sans suivi (n = 140)	Avec suivi (n = 260)	p
DM cliniques	13,7 ± 17,4	4,9 ± 13,4	< 0,0001
DM paracliniques	25,5 ± 25,6	10,3 ± 18,9	< 0,0001
DM traitements	19,5 ± 21,2	16,0 ± 27,2	0,26
DM EI/EIG	—	6,42 ± 19,6	—
DM autres	6,1 ± 17,0	4,5 ± 15,2	0,42
DM totales	16,7 ± 15,4	11,8 ± 16,0	0,003

EI/EIG : évènements indésirables/évènements indésirables graves.

Tableau 3 Pourcentage de données manquantes (DM) avec suivi et sans suivi en fonction de la présence ou non d'un attaché de recherche clinique (ARC).

Types variables (%)	Sans ARC (n = 200)			Avec ARC (n = 200)		
	Sans suivi (n = 120)	Avec suivi (n = 80)	Total	Sans suivi (n = 20)	Avec suivi (n = 180)	Total
DM cliniques	15,0 ± 18,3	9,6 ± 17,8	12,8 ± 18,3	6,2 ± 6,7	2,8 ± 10,2	3,2 ± 9,8
DM paracliniques	25,5 ± 25,6	19,2 ± 18,5	23,5 ± 23,7	—	7,8 ± 17,9	7,8 ± 17,9
DM traitements	19,5 ± 21,2	37,8 ± 33,2	26,9 ± 28,1	—	5,4 ± 14,9	5,4 ± 14,9
DM EI/EIG	—	22,7 ± 35,9	22,7 ± 35,9	—	2,24 ± 8,94	2,24 ± 8,94
DM autres	7,3 ± 18,4	6,6 ± 17,0	7,1 ± 17,9	0	3,1 ± 13,6	2,5 ± 12,3
DM totales	18,7 ± 15,7	27,2 ± 17,6	22,1 ± 17,0	4,3 ± 3,8	4,9 ± 8,8	4,8 ± 8,4

EI/EIG : événements indésirables/événements indésirables graves.

**Figure 1.** Graphique d'interaction entre le type d'étude (avec suivi et sans suivi) et la présence ou non d'un attaché de recherche clinique (ARC) en terme de pourcentage de données manquantes totales (DM).

points forts est que nous présentons un effectif important et que nous avons utilisé des méthodes statistiques rigoureuses. Néanmoins, certaines limites méritent d'être soulevées. Le caractère monocentrique de notre étude ne permet pas de garantir la généralisabilité des résultats. De plus, une analyse qualitative aurait certainement été un plus.

Conclusion

Si la délégation du recueil des données à un personnel dédié permet de faciliter la conduite des protocoles de recherche en libérant du temps investigateur, il est démontré qu'elle permet de diminuer significativement la quantité

de données manquantes. Les ARC/TEC/SC disposent d'un temps dédié pour veiller au meilleur remplissage du cahier d'observation à partir des documents sources. Ils peuvent assister l'investigateur dans la conduite des visites dans le but de veiller à la présence des données dans les documents sources. Ils peuvent également effectuer une recherche de données réintégrables aux données sources afin d'en optimiser le contenu. Par ces caractéristiques, les ARC/TEC/SC permettent d'optimiser quantitativement le recueil de données et donc d'améliorer la qualité globale. Cette optimisation se démontre sur toutes les catégories de données : cliniques, para-cliniques, celles relatives aux traitements, aux événements indésirables graves ou non, ainsi que les informations particulières propres à chaque cahier d'observation.

Déclaration de liens d'intérêts

Les auteurs déclarent ne pas avoir de liens d'intérêts.

Références

- [1] National Research Council. The prevention and treatment of missing data in clinical trials. Washington DC: National Academies Press; 2010. <https://www.nap.edu/catalog/12955/the-prevention-and-treatment-of-missing-data-in-clinical-trials>, [Consulté le 6 novembre 2017].
- [2] European Medicines Agency. Guideline on missing data in confirmatory clinical trials; 2010. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/09/WC500096793.pdf, [Consulté le 6 novembre 2017].
- [3] Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
- [4] Singhal R, Rana R. Intricacy of missing data in clinical trials: deterrence and management. *Int J Appl Basic Med Res* 2014;(1): S2–5.