Modélisation du système de gestion de l'information lors des essais cliniques : Une approche par la dynamique des systèmes

Céline Bérard,
École des Sciences de la Gestion,
Université du Québec à Montréal
berard.celine@courrier.uqam.ca

L. Martin Cloutier,
Groupe de recherche en MAnagement et Transformation
des Entreprises de Biotechnologie (MATEB),
Département de Management et Technologie
École des Sciences de la Gestion,
Université du Québec à Montréal
cloutier.martin@uqam.ca

Luc Cassivi,
Groupe de recherche en MAnagement et Transformation
des Entreprises de Biotechnologie (MATEB),
Département de Management et Technologie
École des Sciences de la Gestion,
Université du Québec à Montréal
cassivi.luc@uqam.ca

ISSN 1718-1534

Résumé

Cet article s'intéresse à l'évaluation de la performance des systèmes d'information gérant les essais cliniques. La complexité et les spécificités du secteur biopharmaceutique soulèvent en effet la problématique de gestion de l'information dans l'exécution de ce processus. Les objectifs principaux sont de mettre en évidence les impacts d'une automatisation du processus de traitement de données sur les projets de gestion de l'information lors des essais cliniques. Plus précisément, la performance des systèmes d'information est évaluée en termes de coût, qualité et délai. L'approche privilégiée est l'utilisation de la dynamique des systèmes, approche enrichie par l'utilisation de cadres théoriques complémentaires. Une étude de cas au sein d'une CRO a été réalisée, afin d'élaborer un modèle de simulation par la dynamique des systèmes. Ce modèle a par la suite permis d'effectuer une analyse comparative de différents systèmes d'information et entre autres, d'anticiper les conséquences de la variation des taux d'erreur lors de la transmission des données. Les résultats observés ont notamment mis en évidence qu'une automatisation performante du processus favorise l'exécution de multi-projets relatifs aux essais cliniques

Mots Clés : essais cliniques, gestion de l'information, systèmes d'information, gestion de projets, mesure de la performance

Introduction

L'industrie pharmaceutique est l'une des plus grandes industries manufacturières dans le monde avec des ventes de produits estimées à environ 300 milliards $US par an. Cette industrie est celle qui a les activités de R-D les plus intenses (Santos 2003). Elle est dédiée à la recherche, au développement et à la commercialisation de médicaments. Le processus complet de mise sur le marché d'un produit est particulièrement long, complexe, fortement réglementé et coûteux.

Les essais cliniques, qui s'inscrivent dans l'étape de développement de nouveaux médicaments, consistent à tester les médicaments, afin de décider ou non de leur commercialisation. Plus précisément, les essais cliniques nécessitent quatre phases. Lors de la phase I, doivent être réalisés des tests sur un nombre important de personnes non malades et volontaires. Puis la phase II consiste à évaluer l'efficacité et les effets secondaires du médicament sur des patients volontaires. Cette évaluation est ensuite vérifiée sur un nombre plus important de patients lors de la phase III. Enfin, la phase IV inclut l'approbation par l'agence régulatrice et le suivi du médicament (Alshawi 2003). Une gestion efficace de l'information lors des essais cliniques est primordiale pour l'industrie pharmaceutique. Les entreprises de ce secteur — cherchant, par des moyens technologiques, à diminuer leurs coûts de R-D de nouveaux médicaments et à réduire la durée de ces activités (Augen 2002) — se tournent alors vers les systèmes d'information, afin de supporter les flux informationnels qu'elles doivent traiter lors de ces quatre phases.

De nombreux logiciels sont ainsi aujourd'hui disponibles pour faciliter le traitement de l'information liée aux essais cliniques. Cependant, ces logiciels n'offrent pas tous les mêmes fonctionnalités, et le niveau de performance de ces fonctionnalités peut, de plus, varier. En outre, le besoin de

chacune des entreprises en matière de gestion de l'information diffère. À titre d'exemple peuvent être citées la culture organisationnelle (certaines privilégient la rapidité du traitement de l'information, d'autres favorisent la probabilité d'erreur nulle...), ou encore les ressources financières (une *Contract Research Organisation* ne peut investir autant qu'un grand groupe pharmaceutique dans son parc technologique et logiciel). Dans ce contexte, le problème de sélection d'un système d'information approprié aux critères spécifiques du processus des essais cliniques mais également aux politiques de l'entreprise se pose.

Ainsi, la principale question à laquelle s'intéresse cet article est la suivante : comment les systèmes d'information améliorent-ils l'exécution et la gestion du processus des essais cliniques, en termes de délai, coût et qualité ? Autrement dit, l'objectif est de mettre en évidence les impacts d'une automatisation du processus sur les projets de gestion de l'information relative aux essais cliniques. Dans cette perspective, cet article est organisé comme suit. Tout d'abord, un rapide survol de l'utilisation des technologies de l'information (TI) dans le domaine de la santé est présenté. Puis, le design de la recherche, ainsi que ses étapes, sont expliqués. Ceci permet ensuite d'appliquer la dynamique des systèmes et de présenter les concepts et modèles d'évaluation élaborés. Enfin, avant de conclure sur les contributions et limites de cette recherche, certains résultats sont analysés.

Les TI et le domaine de la santé

L'émergence des technologies de l'information est l'une des plus grandes évolutions technologiques de ces dernières années et il apparaît clairement que le secteur pharmaceutique n'est pas en reste quant à son utilisation. De manière générale, les TI sont aujourd'hui couramment utilisées dans le domaine de la santé. Notamment, Anderson (2002) s'est intéressé à l'apport des

systèmes d'information dans les services médicaux et aux impacts de ces technologies sur l'environnement organisationnel. Les systèmes d'information concernés permettent par exemple de gérer les prises de décisions médicales en milieu hospitalier, ou encore de détecter et limiter les erreurs médicales. Concernant plus spécifiquement le développement de nouveaux médicaments, Augen (2002) montre dans une de ses études que les technologies de l'information jouent un rôle de plus en plus important dans le pipeline de développement de produits pharmaceutiques. Depuis les 25 dernières années, le processus de développement de médicaments et la variété des TI disponibles ont co-évolué à un tel point qu'ils sont devenus des composantes indissociables du pipeline de R-D. Ainsi, l'apport que les TI ont sur le processus de la découverte de médicaments est devenu un facteur crucial et stratégique pour les entreprises.

Ce constat se vérifie également dans le contexte spécifique des essais cliniques, étant donné que ceux-ci cumulent de grandes quantités d'information et qu'une gestion efficace de l'information lors des essais cliniques s'avère être primordiale. En effet, l'importance de la qualité de l'information transitant lors des expérimentations cliniques ne peut être sous-estimée (Strause et Vogel 1999), ayant des répercussions sur la santé elle-même des patients sous médicamentation. Outre l'importance de la qualité, un autre aspect est également à considérer, soit celui du non respect des échéances, dont souffre fréquemment l'exécution des essais cliniques : 44 % des projets ont entre un et six mois de retard et 12 % des projets ont plus de six mois de retard (Rowe et al. 2002). L'impact négatif de ces retards est facilement imaginable sur le processus total de R-D de nouveaux médicaments : utiliser une mauvaise information pourrait conduire à des conséquences dramatiques sur la santé des individus. De plus, la réglementation est devenue plus stricte au regard de ces étapes : depuis la dernière décennie, le nombre de patients requis pour les essais cliniques a quasiment doublé

(Rowe et al. 2002), augmentant de manière considérable la quantité d'information à traiter. Ces constats justifient le besoin d'automatiser les processus de traitement de l'information, afin de diminuer les délais et coûts associés, mais également de réduire le risque d'informations biaisées, autrement dit, d'améliorer la qualité des données.

Néanmoins, le système de traitement de l'information lors des essais cliniques se révèle complexe. Comme le soulignent Glouberman et Zimmerman (2002), le système de la santé est un système complexe, et qui doit donc être traité comme tel. De ce fait, les problématiques relatives à ce domaine sont de plus en plus appréhendées par des approches telles que la dynamique des systèmes (McDaniel et Driebe 2001). Il s'agit d'une méthode interdisciplinaire, fondée sur la théorie des dynamiques non-linéaires et sur l'identification des rétroactions dans les systèmes, permettant d'améliorer la compréhension et la gestion de systèmes complexes (Sterman 2000). Plus précisément, ces principes offrent deux modes de modélisation, soit le mode qualitatif, basé sur l'utilisation de diagrammes d'influence, et le mode quantitatif, caractérisé par des modèles de simulation systémique (Sterman 2000). Plusieurs auteurs ont mis en évidence les nombreux avantages qu'apporte la simulation systémique, méthode quantitative sous-jacente à cette approche. Anderson (2002) précise notamment qu'avec ces calculs, les modifications de systèmes dans le domaine médical et de la santé, ainsi que l'amélioration de leurs processus, peuvent être étudiées et mesurées. Les comportements futurs du système peuvent donc être anticipés, et ce, sans aucune perturbation sur les comportements actuels du système. Selon Wiendahl et al. (2003, 30), la simulation sur ordinateur « *offre la possibilité de décrire et d'analyser le comportement de systèmes existants dans le cas de modification des conditions et des différents paramètres* ».

La dynamique des systèmes constitue donc un ensemble de méthodes pouvant supporter l'évaluation de la performance des systèmes d'information (Wolstenholme et al. 1993), et est particulièrement utile et adaptée dans l'étude de processus complexes dans le domaine de la santé publique notamment (Homer et Hirsch 2006; Sterman 2006). De ce fait, cette approche a été privilégiée pour répondre à la question principale de la recherche. Il s'agit de produire un modèle générique pouvant simuler le processus de gestion des essais cliniques, sous diverses hypothèses de délai, de qualité et de coût reliées au processus. La méthode de recherche associée est détaillée ci-après.

Méthode de recherche

L'hypothèse de travail qui a conduit cette recherche est qu'il existe une dynamique reliée au temps et aux interrelations qui est d'ordre « générique » dans le processus de gestion de l'information lors des essais cliniques. Le terme « générique » signifie que la structure du processus caractérisée par le modèle est identique, peu importe le contexte de son application. Dans le cas de figure examiné, ce processus est strictement réglementé. En effet, toutes les entreprises qui gèrent ce processus sont sujettes à la même contrainte de gestion de l'information et peuvent trouver des moyens technologiques de gérer l'information dans les produits logiciels actuellement disponibles commercialement. Du fait de cette nature générique, la méthode de recherche favorisée a été la réalisation d'une étude de cas dans une CRO (*Contract Research Organization*), sous-traitant majeur de l'entreprise pharmaceutique. Ce choix a en effet été motivé par le fait que, depuis les années 1980, les entreprises pharmaceutiques impartissent de plus en plus les aspects opérationnels des tests cliniques à ces sociétés de recherche clinique. De ce fait, le nombre de contrats d'impartition a augmenté à une vive allure (Piachaud 2002, Azoulay 2003).

Externaliser une partie des activités permet notamment aux entreprises pharmaceutiques de se recentrer sur leurs compétences de base, d'atteindre un certain niveau de flexibilité et d'efficacité dans leurs processus de R-D, d'améliorer leurs capacités internes et aussi d'avoir un meilleur accès à la technologie, aux connaissances et aux compétences (Piachaud 2002).

Étant donné la nature générique du processus de gestion de l'information lors des essais cliniques, étendre l'étude à plus d'une CRO n'a pas été nécessaire pour la compréhension du processus. Ceci fut mis en évidence par les interlocuteurs de la CRO mais également par une analyse des logiciels. Cette méthode de recherche est typiquement employée dans l'exécution de recherche de cas en dynamique des systèmes (Sterman 2000). Cette présente étude s'est apparentée à une véritable immersion dans le quotidien des employés de la CRO afin de maîtriser leur méthode de travail, leurs besoins et leurs attentes, et d'obtenir l'information nécessaire à la compréhension de la gestion de l'information dans un tel contexte. Les trois principaux interlocuteurs ayant participé à cette collecte de données et d'expertise ont été le responsable de la gestion des affaires, le responsable de l'équipe informatique, et enfin, l'expert finance. De plus, la collecte de données s'est majoritairement basée sur quatre projets précédemment réalisés, ceux-ci concernant des essais cliniques de phase IV. Le terrain de recherche était réellement approprié au contexte d'étude, du fait que cette entreprise se heurtait à la problématique d'investissement TI. En effet, le niveau d'automatisation du processus restait relativement faible, et la CRO avait atteint sa limite de croissance, par manque de ressources disponibles.

Une deuxième source d'information s'est avérée nécessaire dans ce projet de recherche : elle concernait les fournisseurs de logiciels. Dans ce cas, une revue de la littérature et l'étude de

documentations fournies par les éditeurs de logiciels ont suffi à faire un état des lieux du marché. Enfin, pour l'obtention de données plus détaillées, liées aux différents logiciels et à leurs différentes fonctionnalités, des contacts directs ou encore des demandes de démonstration – pratique courante dans ce domaine – ont été établis. Parmi les logiciels permettant l'automatisation de la gestion de l'information dans ce contexte, peuvent être cités les suivants : Oracle Clinical, Datafax, ClinTrial, ou encore Teleform.

L'étude de cas et la familiarisation avec les systèmes d'information concernant la gestion des essais cliniques a permis l'élaboration d'un modèle de simulation. Étant donnée la complexité du contexte et du processus concernés par cette recherche, un cadre organisant rigoureusement la méthode s'est avéré utile (voir figure 1). Son objectif, outre de guider les étapes de la recherche, était de stocker et présenter l'information requise pour l'élaboration du modèle, et ce, de manière structurée et afin de n'omettre aucun paramètre fondamental dans ce projet de recherche. Ce cadre a principalement été basé sur le processus itératif de modélisation par la dynamique des systèmes proposé par Sterman (2000).

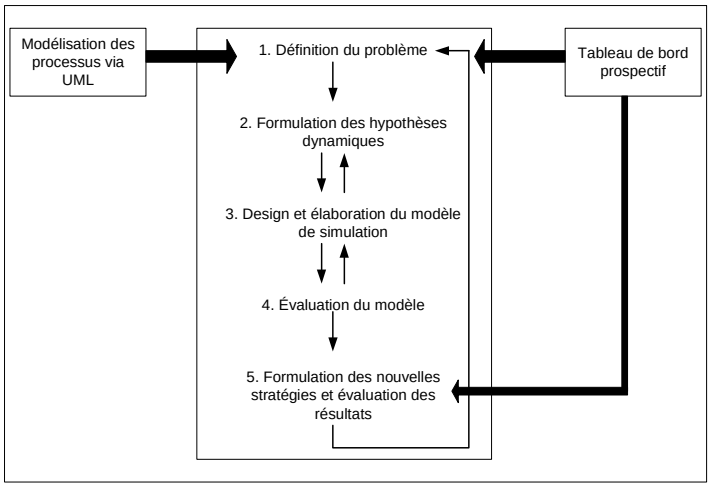

Figure 1 – Cadre conceptuel

Ainsi, dans un premier temps, il a été nécessaire de définir le problème, ainsi que l'objectif du modèle. Pour ce faire, deux méthodes se sont imposées : la modélisation par *Unified Modeling Language* (UML), permettant de représenter le processus d'affaires en tant que tel et le tableau de bord prospectif, proposé par Kaplan et Norton (2001), avec sa carte stratégique et les indicateurs de performance qui doivent être mesurés pour chacun des scénarios. La notation UML permet en effet de modéliser des processus d'affaires, et, plus précisément, le diagramme d'activités (*Version 1.5, OMG Unified Modeling Language Specification, 2003*) sert à modéliser les flux de données, de contrôle, et, de plus, fait intervenir les acteurs. Il permet ainsi de représenter les aspects dynamiques à un niveau assez général. Concernant le tableau de bord prospectif, celui s'intéresse à la stratégie globale d'une entreprise. Il permet de voir les répercussions d'une automatisation d'un processus d'affaires selon quatre axes, à savoir l'axe financier, processus internes, apprentissage et développement, et enfin, client (Kaplan

et Norton 2001). Cependant, il faut noter que l'unité d'analyse du modèle de simulation est le projet de gestion de l'information, et non l'organisation. De ce fait, certains indicateurs n'ont pas été gérés par le modèle de simulation, celui-ci ayant un périmètre d'action plus limité que celui du tableau de bord.

Dans un deuxième temps, les hypothèses dynamiques ont dû être formulées, autrement dit, des diagrammes d'influence ont été élaborés, afin de mettre en évidence les principales rétroactions qui caractérisent la structure du système.

Troisièmement, la formulation du modèle de simulation, ainsi que sa quantification et son calibrage, a été réalisée. Le modèle a été caractérisé en termes de variables d'état et de variables de contrôle, autrement dit de type continu. La représentation de ce modèle a été faite via l'utilisation d'un logiciel informatique, soit Powersim.

Dans un quatrième temps, le modèle a été testé, afin de vérifier si le modèle simulé de référence se rapprochait du comportement réel du système. Cette étape a permis de valider le modèle de simulation et de s'assurer de sa crédibilité.

Enfin, la cinquième et dernière étape a nécessité l'identification de scénarios, c'est-à-dire de nouvelles stratégies éventuelles, et l'analyse des résultats engendrés par le modèle. Plus précisément, les indicateurs de performance du tableau de bord prospectif ont permis de mettre en évidence certains des critères à évaluer lors de cette dernière étape. En d'autres termes, chacun des critères d'évaluation s'apparente à l'un des quatre axes définis précédemment. Par exemple, les critères de coût concernent principalement l'axe financier du tableau de bord, alors que les critères de délai sont majoritairement en lien avec l'axe des processus internes.

Élaboration des concepts et modèles d'évaluation

L'étape ultime de la démarche de recherche pour évaluer la performance des systèmes d'information dans la gestion de l'information lors des essais cliniques a résulté en l'élaboration d'un modèle de simulation. Cette section présente les concepts et modèles élaborés, en lien avec le cadre conceptuel défini précédemment : le diagramme d'activités de la notation UML, le diagramme d'influence et la structure du modèle de simulation. Le tableau de bord n'est pas présenté dans cet article, étant donné qu'il représente la stratégie au niveau organisationnel et non au niveau du projet de gestion de l'information lui-même.

Succinctement, le processus de gestion de l'information lors des essais cliniques s'apparente à un échange d'information entre l'entreprise responsable des essais cliniques, les médecins examinant les patients et l'agence régulatrice. Les entreprises en charge des essais cliniques ont alors pour mission de recevoir les données relatives aux tests cliniques, de les traiter et enfin de les analyser. Il est important de préciser que ces données peuvent être collectées directement en interne, par le biais d'un centre d'appels téléphoniques, ou peuvent provenir de médecins externes à l'entreprise. À noter toutefois que l'information à traiter est plus fréquemment de nature externe et que les processus ne diffèrent que très peu d'une phase à l'autre. Ce processus est représenté via la notation UML dans la figure 2.

La première étape importante de ce processus est la définition du protocole. Il s'agit d'un document déterminant les bases de l'étude à réaliser, tels que le nombre de médecins à recruter, le nombre de patients à enrôler, le nombre de visites à effectuer par patient, le contenu des questionnaires à renseigner, etc. Dès que le recrutement des médecins est effectué et que les formulaires (questionnaires) sont prêts à être distribués aux médecins, l'étude peut débuter. Ces formulaires, communément

appelés « CRF » (*Case Report Form*), sont ensuite à remplir par les médecins recrutés suite à la consultation d'un patient, plus souvent nommée « visite ». Ils peuvent alors s'enquérir d'informations telles que l'identification du patient, les résultats d'analyse sanguine ou autres, le niveau de confort de vie du patient, etc. Le médecin peut ainsi transmettre le ou les formulaires avec les informations requises. Ceci se fait généralement sous forme de télécopie. Les données faxées sont ensuite traitées, autrement dit vérifiées, éventuellement corrigées et enfin, stockées. Dans le cas d'un CRF en erreur (données erronées ou manquantes sur un formulaire), un formulaire listant les informations en erreur est généré et devra être faxé au médecin en vue d'obtenir une correction. Ce formulaire est appelé un « DCF » (*Data Clarification Form*). Lorsque toutes les visites requises ont été effectuées, autrement dit en fin d'étude, les données précédemment stockées sont analysées, et un rapport d'analyse statistique doit être rédigé, afin de guider la décision de l'agence régulatrice.

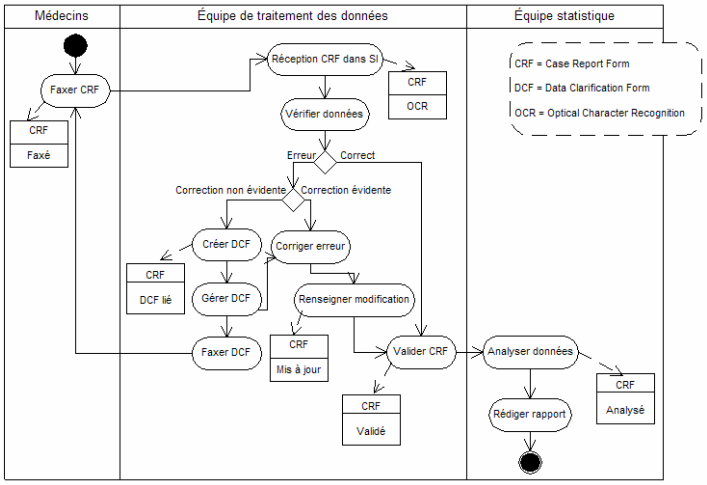

Figure 2 – Diagramme d'activités UML

Ainsi, dans un tel contexte, le rôle principal du système d'information, en plus de supporter la gestion des projets, est la réception des fax, la reconnaissance optique des données (conversion de données « papier » en données « électroniques »), le contrôle des données, la gestion et le suivi des fax ainsi que des erreurs, et enfin le stockage des données en vue de leur consultation et analyse. Par la suite, les principales boucles de rétroaction de ce système ont pu être mises en évidence. La figure 3 illustre le diagramme élaboré et simplifié et il apparaît que les boucles de rétroaction sont principalement de type équilibrage. Plus précisément, sept boucles d'équilibrage (E) et deux boucles de renforcement (R) ont été identifiées[1].

Les boucles E1, E2 et E3 concernent la génération des visites. Celles-ci nécessitent dans un premier temps l'enrôlement des patients, qui est en fonction d'un objectif déterminé en début d'étude : il s'agit de la boucle E1. Toutefois, certains patients précédemment enrôlés sont supprimés par la suite de l'étude, soit en raison d'un abandon du patient lui-même, soit en raison de critères non respectés. Le taux de rejet impacte ainsi négativement le nombre de patients, et ceci est illustré dans la boucle d'équilibrage E2. La boucle E3 s'intéresse à la génération elle-même des visites, qui dépend du nombre de patients à visiter, et qui requiert un certain délai avant d'être complétée. La génération des visites engendre l'envoi par télécopie des CRF. Ceux-ci sont transférés dès réception sur le système

[1] Il existe deux types de boucles de rétroaction pour caractériser la structure d'un système. Une boucle est dite de 'renforcement' si elle perpétue une action de croissance et de décroissance sans contrainte. Une boucle de rétroaction est dite 'd'équilibrage' si elle engendre un sentier dynamique visant l'équilibre d'un système. Les systèmes complexes sont des assemblages de boucles de renforcement et d'équilibrage en interaction. Selon la syntaxe établie par les principes de la dynamique des systèmes (Sterman, 2000), un diagramme d'influence représente la direction d'une influence entre un ensemble de variables par l'utilisation des symboles '+' et '-'. Une boucle de renforcement, représentée par (R) suivie d'un nombre qui l'identifie, ne possède que des influences qui vont dans la même direction entre les variables ou un nombre pair d'influences inverses '-' dans la boucle de rétroaction. Une boucle de rétroaction d'équilibrage est représentée par (E) suivie d'un nombre qui l'identifie. Cette dernière possède un nombre impair de relations opposées entre les influences d'une boucle de rétroaction.

d'information, ce qui correspond à la boucle d'équilibrage E4. Ces CRF peuvent ensuite soit être directement sauvegardés dans la base de données, soit nécessiter la gestion d'un DCF, ou encore, une correction manuelle (traduite dans la boucle E5). Dans le cas où un DCF doit être géré, le CRF initial peut être corrigé manuellement, ou peut requérir un nouveau CRF (faxé de nouveau par le médecin après un certain délai) : il s'agit de la boucle de renforcement R1. Il faut alors noter que certains DCF restent sans réponse de la part du médecin, empêchant le traitement du CRF correspondant. Enfin, les boucles E6 et E7 s'intéressent plus particulièrement au suivi des CRF et à la qualité des données. En effet, des erreurs peuvent survenir lors, par exemple, de l'assimilation du statut d'un CRF (E6), ou encore, des données peuvent être manquantes ou erronées dans la base de données elle-même (E7).

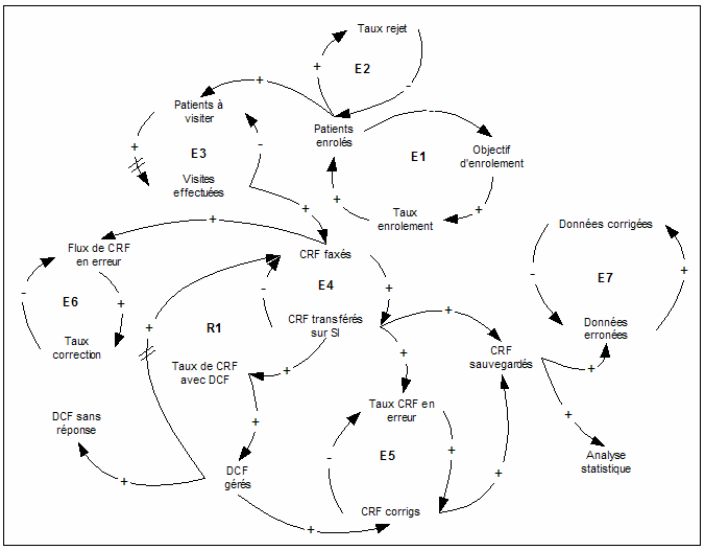

Figure 3 – Diagramme d'influence

L'identification de ces boucles de rétroaction a alors servi de base à l'élaboration du modèle quantitatif, autrement dit, du modèle de simulation. La collecte de données (au sein de la CRO

et auprès des fournisseurs de logiciels) a permis l'adaptation de ce modèle, avec la quantification des variables du modèle. Cette quantification a de plus permis le calibrage et le test du modèle dit de référence. La figure 4 représente la structure générale du modèle de simulation élaboré.

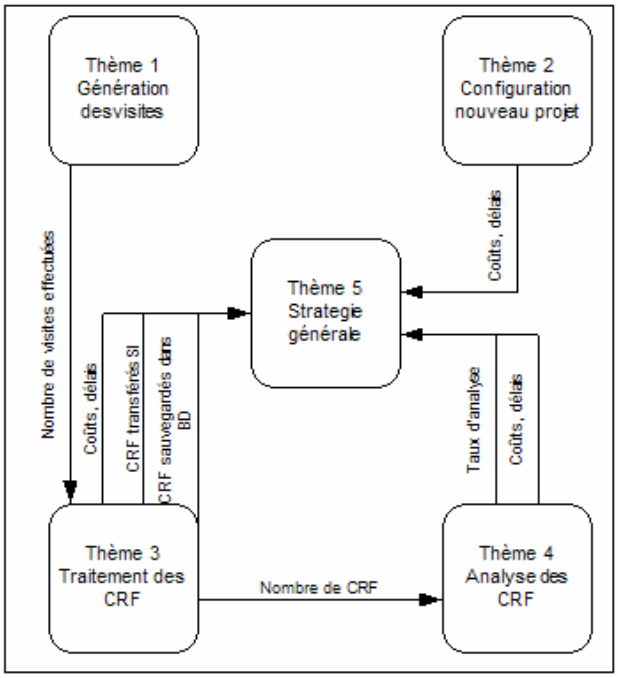

Figure 4 – Structure générale du modèle de simulation

Il apparaît ainsi que le processus a été analysé selon cinq thèmes, donnant naissance à cinq sous-secteurs de modèles de niveaux-taux, et mettant en évidence les variables interagissant entre ces thèmes :

1) La génération de visites par les médecins, liée aux caractéristiques de l'étude à réaliser, qui sont principalement liées au volume de l'étude ;

2) Le lancement de l'étude, autrement dit les processus liés à la configuration requise à chaque nouveau projet, qui prend en considération toutes les étapes de préparation, à réaliser avant la gestion de l'information en tant que telle ;
3) Le traitement des CRF, donc des données, c'est-à-dire le cœur de la gestion de l'information lors du processus des essais cliniques ;
4) L'analyse statistique des CRF ;
5) La stratégie générale de l'entreprise qui est présente tout au long de l'étude et oriente inévitablement les différentes actions.

Résultats

La démarche de recherche, telle que décrite dans la section précédente et qui s'est inspirée des principes méthodologiques documentés par Sterman (2000), a permis la structuration et l'élaboration des différents moyens d'analyses, soit la formulation de nouvelles stratégies et l'évaluation des résultats. Le modèle de simulation élaboré a permis de mettre en évidence les principaux impacts d'une automatisation plus ou moins poussée du processus de gestion de l'information lors des essais cliniques, en procédant à une analyse comparative de différents systèmes d'information. De plus, différents scénarios ont été simulés, comme notamment la variation du volume d'affaires, du taux d'erreur relatif à la performance de la reconnaissance optique du système d'information, ou encore, du nombre de ressources humaines composant l'équipe TI. L'analyse des résultats a principalement été axée sur les indicateurs de performance (en termes de délai, coût et qualité), tels qu'identifiés dans le tableau de bord prospectif. Cette section présente tout d'abord les résultats engendrés par l'analyse comparative de différents systèmes d'information. Puis, l'accent est mis sur l'étude du scénario traitant de la variation du taux d'erreur lié à la reconnaissance optique des données.

L'étude des logiciels disponibles pour l'analyse a permis de déterminer trois classes de systèmes d'information : les logiciels de petite envergure (λ), de moyenne envergure (σ) et enfin de grande envergure (γ). L'envergure concerne alors à la fois les fonctionnalités proposées par le logiciel, mais également son coût, ou encore la taille et la place sur le marché de l'éditeur. Les logiciels de la classe λ, autrement dit les « petits » logiciels, sont donc ceux offrant le moins de fonctionnalités, et n'automatisent que peu de transactions. Cette classe peut notamment concerner des logiciels tels que Clindex ou encore Registrat. En ce qui a trait au type σ, le logiciel Datafax est le plus représentatif de cette classe dite « intermédiaire ». Ce type de logiciels offre la quasi-totalité des fonctionnalités attendues. Enfin, la classe γ représente des logiciels tels qu'Oracle Clinical ou Clin'Trial, autrement dit, les « géants » parmi les éditeurs de logiciels dans le domaine des essais cliniques. Les logiciels de ces trois catégories ont été comparés sur la base d'un projet « type », s'étalant sur plus d'un an. Plus précisément, et d'après les moyennes calculées sur plusieurs projets réalisés en phase IV par la CRO étudiée, le projet « type » concerne 1000 patients enrôlés, avec trois visites requises par patient. Il a été estimé à 6000 le nombre total de CRF initiaux à traiter, et à 180 000$ le revenu généré.

Les résultats obtenus sont les suivants : en termes de délai, les technologies ont un impact important sur la configuration des CRF ; en termes de qualité, il a été mis en évidence le fort impact du suivi automatisé des fax reçus et envoyés, ainsi que des contrôles qualité des données en base de données ; en termes de coût, les logiciels influencent de manière non négligeable les coûts relatifs à la gestion de projet, et de manière conséquente les coûts liés à la gestion de la qualité et à la configuration en début de projet. Il a également été mis en évidence le fait que le choix d'une technologie ne se répercute que très peu sur le processus du traitement des CRF, et ce, tant en terme de coût que de délai.

Cependant, il est important de préciser qu'une automatisation de la gestion des DCF, engendrée par des solutions de type σ ou γ, permet une meilleure réactivité de la part du médecin lors de la demande de correction.

Il semble alors évident que les solutions de type σ ou γ offrent de réels avantages. L'investissement technologique requis est plus élevé que pour les petits logiciels, mais il peut être amorti avec un volume d'affaires représentant près de cinq projets par année. La différence entre les classes σ et γ semble minime et il serait alors tentant de recommander les solutions de type σ, dont le coût est moins important. Cependant, d'autres facteurs d'ordre qualitatif sont également à prendre en considération. Notamment, les technologies doivent être analysées également en terme de pérennité de solution mais aussi du fournisseur lui-même. Par exemple, il est plus sécurisant d'opter pour un éditeur leader sur le marché, afin de s'assurer que les maintenances, mises à jour, intégration avec d'autres technologies informatiques futures, etc. soient accessibles.

Afin d'approfondir cette analyse, des scénarios ont été testés, dont celui de la variation du taux d'erreur OCR. En effet, l'automatisation de base requise dans le contexte de gestion de l'information lors des essais cliniques requiert la reconnaissance optique des données (OCR) afin de reconnaître informatiquement les informations faxées par les médecins. Cependant, certaines erreurs peuvent être liées à cette reconnaissance optique, ce qui est traduit en tant que taux d'erreur OCR, donc le pourcentage de formulaires faisant l'objet d'un ou plusieurs champs non reconnus correctement. Ce taux d'erreur est principalement lié à la performance du système d'information mis en œuvre mais également à la clarté du formulaire faxé. Par clarté sont sous-entendues la mise en forme elle-même du formulaire (espacement entre les différents champs à renseigner) et la lisibilité de l'écriture du médecin. Initialement

estimé à 15% (situation de référence, scénario 1), le taux d'erreur sera modifié d'une part à 0% (scénario 2), d'autre part à 50% (scénario 3) et enfin à 70% (scénario 4). Ainsi, les impacts sur le processus de la performance du système liée à la reconnaissance optique et de l'effort mis en œuvre en tant que clarté du formulaire ont pu être anticipés lors de l'analyse de ce scénario.

L'analyse des résultats a mis en évidence que la variation du taux d'erreur OCR a directement un impact sur le nombre de CRF à corriger, réduisant ou augmentant, dépendamment de la variation du taux, les activités de traitement des CRF et ainsi les coûts d'exploitation d'un projet. Le tableau 1 identifie les résultats, autrement dit, les données variant en fonction du taux d'erreur.

Tableau 1 – Impacts de la variation du taux d'erreur OCR

Scénario	**1**	**2**	**3**	**4**
Taux d'erreur relatif à OCR	15	0	50	70
CRF à corriger dû à une erreur OCR (nombre)	900	0	3000	4200
Traitement des CRF (nombre de jours)	185	176	207	220
Coûts liés au traitement des CRF ($)	14 815	14 056	16 586	17 597
Marge brute bénéficiaire par projet (% de revenus)	40,9	41,3	40,1	39,6

Il apparaît ainsi qu'un taux d'erreur OCR nul favoriserait une réduction des coûts de 20% relativement à un taux d'erreur de 70 %. Comparativement à la situation de référence, soit avec un taux de 15 %, l'augmentation du taux d'erreur OCR entraînerait une augmentation des coûts liés au traitement des CRF de près de 1770 $ pour le scénario 3 et de près de 2800 $ pour le scénario 4. Ceci est dû à une variation de la durée des activités de traitement des CRF : le nombre de jours associés à la correction manuelle des CRF se verrait augmenter. En effet, tel que présenté dans le tableau 1, le nombre de jours alloué au traitement des CRF évolue en fonction du taux d'erreur. Celui-ci est ainsi de 185 jours pour la situation de référence, et serait réduit à 176 jours avec un taux d'erreur nul, ou atteindrait 207 jours, voire 220 jours, avec un taux respectivement de 50 % et de

70 %. La marge brute bénéficiaire est alors impactée par ce taux d'erreur, mais les variations restent minimes : estimée à 40,9 % dans le cas de référence, elle n'atteint que 41,3 % avec un taux d'erreur nul et n'est réduite qu'à 39,6 % avec un taux d'erreur de 70 %. Cependant, il est vrai que la simulation avec un taux d'erreur nul n'est pas un comportement pouvant être réel. En effet, il est quasiment impossible d'atteindre un taux de réussite de reconnaissance optique des données à 100%.

Il peut alors être supposé que moins le nombre de CRF à corriger est important, plus le projet sera rapidement complété. Cependant, tel n'est pas le cas, comme le montre la figure 4, qui présente le flux des CRF sauvegardés dans la base de données, et donc, prêts à être analysés en fin de processus. Ainsi, malgré une vitesse plus accrue de la sauvegarde des CRF, la totalité des CRF est traitée seulement dans le courant du treizième mois, et ce, quel que soit le scénario. La figure 5 permet de comprendre ce phénomène. En effet, celle-ci met en évidence que les CRF sont corrigés sur la même période, et ce, quel que soit le nombre de CRF concernés. Ainsi, la CRO corrige effectivement plus de CRF dans le cas des scénarios 3 ou 4, mais réalise cette activité sur le même laps de temps que le scénario 1 ou 2. Il faut alors noter que même dans le cas du deuxième scénario, le nombre de CRF à corriger n'est pas nul. Ceci s'explique par le fait que d'autres erreurs, autres que relatives à OCR, peuvent intervenir.

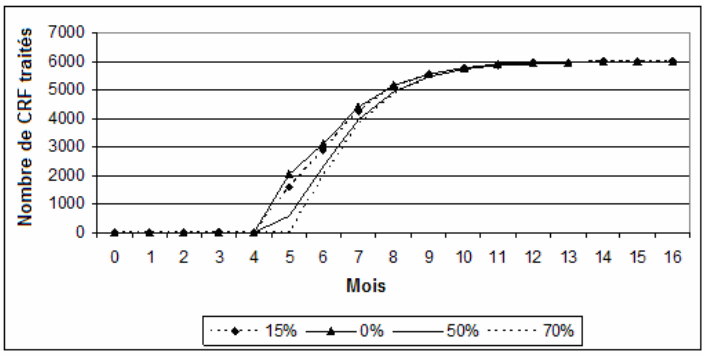

Figure 5 - Impacts de la variation du taux d'erreur OCR sur la durée du projet

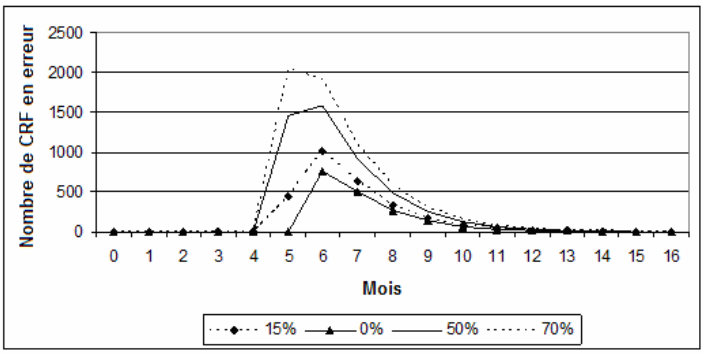

Figure 6 - Impacts de la variation du taux d'erreur OCR sur le nombre de CRF en erreur

En résumé, il est important de préciser que la variation du taux d'erreur OCR n'a quasiment pas d'impact sur la durée d'exécution du projet. En effet, quelle que soit la stratégie adoptée, les délais associés au traitement des CRF, voire du projet dans son intégralité, sont beaucoup plus dépendants de la fréquence et cadence de l'enrôlement des patients et des visites s'ensuivant, que de facteurs endogènes au processus de gestion de l'information. Ainsi, l'augmentation ou la diminution du nombre de CRF à corriger n'a qu'un impact minime sur le

déroulement d'un projet. Il a donc été mis en évidence que les répercussions importantes concernent principalement l'effort en ressources humaines requis. La variation du taux d'erreur OCR et de manière générale, du niveau de performance des logiciels, devrait alors avoir de réels impacts dans le cadre de multi-projets, les ressources pouvant être allouées à d'autres tâches dans le cadre de taux d'erreur faible.

Conclusion

Cet article s'intéresse à l'évaluation de la performance des systèmes d'information en soutien aux projets de gestion de l'information lors des essais cliniques. L'approche privilégiée pour la réalisation de cette recherche a été l'utilisation des méthodes de la dynamique des systèmes, en vue d'élaborer un modèle de simulation. Ce modèle avait alors pour objectifs principaux de mettre en évidence les répercussions d'une automatisation du processus et de favoriser la bonne compréhension des enjeux informationnels auxquels doivent faire face les entreprises en charge des essais cliniques. L'analyse du processus de la gestion de l'information lors des essais cliniques et sa modélisation ont permis de souligner la complexité d'un tel système. En effet, il est caractérisé par de nombreuses variables et relations de cause à effet non-linéaires, en l'occurrence, des boucles de rétroaction de type équilibrage. Ces dynamiques rendent donc difficile la bonne compréhension du comportement actuel et futur du système. Par ailleurs, il a été nécessaire de structurer le processus de modélisation du modèle de simulation, en s'appuyant sur des concepts et cadres tels que la modélisation des processus via UML ou encore le tableau de bord prospectif. Le cadre conceptuel nouvellement créé et suivi apporte alors une réelle contribution quant à la démarche de recherche proposée, tout en respectant les théories avancées par Sterman (2000) quant à l'approche de la dynamique des systèmes. L'application dans ce contexte complexe en est

d'autant plus intéressante, et les avantages que peuvent apporter la simulation sont mis en évidence.

Le modèle ainsi élaboré a permis de réaliser une analyse comparative des technologies logicielles, et, de plus, des scénarios ont pu être testés. Cet article a illustré les impacts de la variation du taux d'erreur lié à la reconnaissance optique des données. L'analyse des résultats a permis de mettre en évidence qu'un système d'information performant favorise l'exécution de multi-projets, et ainsi la croissance du nombre de projets pouvant être entrepris parallèlement, sans toutefois impacter la durée relative à un unique projet. Ce constat est alors fondamental, étant donné que le nombre de médicaments testés simultanément, combiné à la qualité des données relatives à l'expérimentation clinique, a des répercussions sur la santé elle-même des patients en attente de médicamentation appropriée et fiable. La fiabilité des données devient dès lors critique, puisqu'elle constitue un intrant clé dans les analyses statistiques des résultats des essais cliniques. De plus, les résultats des essais cliniques dépendent de la significativité statistique. Ces résultats servent donc de critère d'acceptabilité ou de rejet des molécules à l'étude lors des essais cliniques. La question de la qualité des données qui entrent dans ces calculs statistiques prend donc toute son importance et minimiser les risques associés aux résultats sur lesquels les décisions cliniques reposent se révèle crucial.

Cependant, cette recherche a ses limites : ce projet se limite à l'aspect générique de la gestion de l'information lors des essais cliniques. Malgré ce caractère générique, et comme cela a été précédemment exposé, les besoins et contraintes de gestion des essais cliniques diffèrent selon les entreprises. Appliquer le modèle à d'autres CRO serait ainsi intéressant pour valider le modèle générique et confirmer les conclusions. De même, il semble difficile d'assumer que les autres types d'entreprises du secteur biopharmaceutique soient concernés par cette étude. En

effet, les grands groupes pharmaceutiques ou les petites sociétés spécialisées en biotechnologie ont des caractéristiques différentes des sociétés de recherche clinique, mais n'ont pas été étudiés dans ce projet.

Bibliographie

Alshawi, S., Saez-Pujol, I., Irani, Z. (2003), «Data warehousing in decision support for pharmaceutical R&D supply chain», *International Journal of Information Management*, 23, 259-268.

Anderson, J.G. (2002), «Evaluation in health informatics: Computer simulation», *Computers in Biology and Medicine*, 32, 151-164.

Augen, J. (2002), «The evolving role of information technology in the drug discovery process», *Drug Discovery Today*, 7, 315-323.

Azoulay, P. (2003), Acquiring knowledge within and across firm boundaries: evidence from clinical development. Consulté sur internet: *http://www.nber.org/papers/w10083, le 16/04/2007*.

Glouberman, G., Zimmerman, B. (2002), «Complicated and complex systems: What would successful reform of medicare look like?», *Commission on the Future of Health Care in Canada*, Discussion Paper No 8, 1-37.

Homer, J.B., Hirsch, G.B. (2006), «System dynamics modelling for public health: Background and opportunities», *American Journal of Public Health*, 96, 452-458.

Kaplan, R.S., Norton, D.P. (2001), *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*, Boston, Massachusetts: Harvard Business School Press.

McDaniel, R.R., Driebe, D.J. (2001), «Complexity science and health care management», *Advances in Health Care Management*, 2, 11-36.

Piachaud, B.S. (2002), «Outsourcing in the pharmaceutical manufacturing process: an examination of the CRO experience», *Technovation*, 22, 81-90.

Rowe, J. C., Elling, M. E., Hazlewood, J. G., Zakhary, R. (2002), «A cure for clinical trials», *The McKinsey Quartely*, 2, 1-6.

Santos, F.M. (2003), «The coevolution of firms and their knowledge environment: insights from the pharmaceutical industry», *Technological Forecasting & Social Change*, 70, 687-715.

Sterman, J.D. (2000), *Business Dynamics: Systems Thinking and Modeling for a Complex World*, New York, NY, Irwin-McGraw-Hill.

Sterman, J.D. (2006), «Learning from evidence in a complex world», *American Journal of Public Health*, 9, 505-514.

Strause, L.G., Vogel, J.R. (1999), «The clinical research triad: How can we ensure quality in out-sourced clinical trials?», *Quality Management in Health Care*, 7, 23-29.

Wiendahl, H.P., Worbs, J. (2003), «Simulation based analysis of complex production systems with methods of non-linear dynamics», *Journal of Materials Processing Technology*, 139, 28-34.

Wolstenholme, E., Henderson, S. & Gavine, A. (1993), *The Evaluation of Management Information Systems: A Systemic and Holistic Approach*, New York, NY, Wiley & Sons.