Titre: Title:	Optimisation des horaires et de rendez-vous de patients : applications en imagerie médicale et radiothérapie
Auteur: Author:	Dina Ben Tayeb
Date:	2022
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Ben Tayeb, D. (2022). Optimisation des horaires et de rendez-vous de patients : applications en imagerie médicale et radiothérapie [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/10773/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/10773/
Directeurs de recherche: Advisors:	Louis-Martin Rousseau, & Nadia Lahrichi
Programme: Program:	Doctorat en génie industriel

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Optimisation des horaires et de rendez-vous de patients : applications en imagerie médicale et radiothérapie

DINA BEN TAYEB

Département de mathématiques et de génie industriel

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*

Génie industriel

Décembre 2022

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Cette thèse intitulée :

Optimisation des horaires et de rendez-vous de patients : applications en imagerie médicale et radiothérapie

présentée par **Dina BEN TAYEB**

en vue de l'obtention du diplôme de *Philosophiæ Doctor*

a été dûment acceptée par le jury d'examen constitué de :

Michel GENDREAU, président

Louis-Martin ROUSSEAU, membre et directeur de recherche

Nadia LAHRICHI, membre et codirectrice de recherche

François SOUMIS, membre

Angel RUIZ, membre externe

DÉDICACE

À mes parents, et à Chahid

REMERCIEMENTS

Je remercie très sincèrement mes directeurs de recherche Louis-Martin Rousseau et Nadia Lahrichi pour leur professionnalisme, leur temps, et leur patience. Je suis très reconnaissante de leur soutien moral et de leur compréhension, grâce à eux, j'ai pu surmonter des moments difficiles de ma vie et faire l'équilibre entre les études et la famille. Honnêtement, j'avais la chance et l'honneur d'être parmi leur groupe de recherche.

Je voudrais remercier les membres de mon jury pour avoir accepté d'évaluer mes travaux de thèse : Michel Gendreau, François Soumis, et Angel Ruiz.

J'aimerais remercier tous nos collaborateurs du Centre Intégré de Cancérologie de Laval et du Centre hospitalier de l'Université de Montréal, en particulier, Marie-Andrée Fortin, François Guilbert, Amanda Jean-Gilles, Nadine Beauchemin, et Line Carpentier, pour leur temps et leur confiance. Ils m'ont permis de connaitre les défis du monde de travail durant ma recherche.

Merci aux organismes qui ont participé au financement de mes projets de doctorat : le CRSNG, MITACS, et Elekta. Je remercie également les membres de MAGI, SEP, GERAD, et CIRRELT, particulièrement Khalid Laaziri, pour leur aide.

Je remercie mes amis qui étaient pour moi une deuxième famille, nous avons partagé ensemble des moments de joie et de chagrin.

Merci à mon frère avec qui m'a supporté durant les moments les plus difficiles.

Un grand merci à mes parents pour leur sacrifice et leur soutien. Une mention toute particulière à ma mère qui m'a accompagné durant toutes mes années de doctorat, elle a pris soin de moi et de ma petite famille avec un grand amour. Sa présence m'a donné la confiance et le courage pour continuer tout au long de mon parcours scolaire.

Je remercie infiniment mon cher époux Chahid. Il m'a motivé pour poursuivre mes études. Grâce à son aide, son soutien et ses conseils, j'ai pu avancer dans cette thèse.

RÉSUMÉ

La planification efficiente des ressources humaines et matérielles en santé assure la satisfaction des patients et des centres de santé en maximisant l'accès aux soins et en minimisant les coûts. Or, les administrateurs choisissent souvent des solutions faciles à utiliser ou à implémenter au lieu de celles qui seraient optimales et donc plus efficaces. Cette décision revient à leur manque de connaissance ou bien leur impossibilité d'investir pour adopter de nouvelles approches. Cette thèse exploite des techniques d'apprentissage machine et de recherche opérationnelle pour une meilleure gestion des ressources dans un centre d'imagerie médicale et un centre de radiothérapie. Nous nous focalisons sur des applications réelles en planification des rendez-vous des patients et des horaires des technologues. Nous avons analysé des vraies données grâce aux collaborations réalisées avec le Centre Intégré de Cancérologie de Laval (CICL) et le Centre hospitalier de l'Université de Montréal (CHUM).

L'hétérogénéité des patients et leur différence en termes de catégorie de traitement, mobilité, type de visite, etc. influencent le temps de service au sein des centres de santé. L'assignation de la même durée de traitement (ou encore bloc) à tous les patients engendre donc des coûts liés à la sous-utilisation des ressources, l'attente, et le temps supplémentaire. Le premier article de cette thèse consiste à proposer une approche sans bloc pour la conception des grilles des rendez-vous des patients en radiothérapie. Nous utilisons les données réelles fournies par le CICL. Nous développons en premier lieu un modèle de prédiction à l'aide des techniques d'apprentissage machine : modèle linéaire général, Multivariate Adaptive Regression Splines (MARS), les réseaux de neurones artificiels, l'arbre de classification et de régression. Le meilleur modèle est élaboré par l'arbre de classification et de régression avec une précision de 84%. En se basant sur les valeurs prédites, nous redéfinissons les grilles des rendez-vous, qui sont évaluées à l'aide les règles de séquence et de gestion des patients. Les meilleures solutions sont sélectionnées en fonction du nombre de patients traités, le temps d'attente et le temps supplémentaire. En appliquant la nouvelle approche, nous pouvons atteindre une augmentation de 30% des patients par journée.

Le deuxième article de cette thèse traite le problème de la planification des rendez-vous des

patients et des horaires des technologues. Dans la plupart des centres de santé, ces deux éléments sont traités séparément. Dans notre étude, nous développons deux versions d'un modèle d'optimisation : séquentiel et intégré. La première version consiste à élaborer les horaires des technologues initialement et concevoir les grilles des rendez-vous par la suite; et, la deuxième version combine les deux éléments. Nous visons à assurer une allocation optimale du personnel tout en maximisant l'utilisation des machines, le nombre des patients traités, et la stabilité des horaires et des grilles. Notre approche est évaluée en se basant sur les données réelles du département d'imagerie par résonance magnétique (IRM) du CHUM, en modifiant les règles de travail des technologues et les méthodes de construction des plannings des technologues. Bien que la version séquentielle soit la plus simple, elle donne de bons résultats selon certains indicateurs (comme la stabilité des horaires du technologue), au détriment d'autres indicateurs (comme la stabilité des grilles de rendez-vous), contrairement à la méthode intégrée qui considère tous les facteurs simultanément. En menant une analyse qui clarifie l'impact de chaque solution versus la complexité de l'implantation, nous guidons les administrateurs vers les meilleures solutions et au plan optimal de leur mise en place.

La dernière contribution de cette thèse concerne la replanification des patients en radiothérapie en cas de panne de machine. Nous développons un modèle d'optimisation et une heuristique. Ils sont appliqués dans un cadre bien défini comportant toutes les séquences de priorité des décisions de replanification qui sont : retarder un patient, surréserver un patient, utiliser le temps supplémentaire pour céduler un patient. Nous nous sommes basés dans notre étude de cas sur les données historiques du CICL concernant les temps de traitement, le nombre de séances des traitements et les dates d'échéances des plans de traitement. Vu le manque de documentation liée aux pannes des machines, l'approche proposée est évaluée à l'aide des scénarios générés en modifiant la durée de panne et le taux de remplissage du planning. Les solutions sont comparées par : le nombre de patients retardés, le temps d'attente, et le temps supplémentaire. Nous présentons une analyse détaillée des résultats qui permet aux gestionnaires de comprendre l'impact des décisions prises sur la performance.

ABSTRACT

Efficient planning of human and material health resources ensures the satisfaction of patients and health centers by maximizing access to care and minimizing costs. However, administrators sometimes choose easy-to-use or easy-to-implement solutions instead of optimal and more efficient ones. This decision results in their lack of knowledge or their inability to invest in adopting new approaches. This thesis exploits machine learning and operations research techniques for elaborating patient appointments grids and technologist schedules. We have broached real applications by analyzing real data thanks to the collaborations carried out with the Centre Intégré de Cancérologie de Laval (CICL) and the Centre hospitalier de l'Université de Montréal (CHUM).

The heterogeneity of patients and their difference in terms of treatment category, mobility, type of visit, etc. influence service time in health centres. Therefore, assigning the same duration of treatment to all patients generates costs related to underutilization of resources, waiting time, and overtime. This thesis consists in proposing a Non-block approach for the design of patient appointment grids in radiotherapy. We use the actual data provided by the CICL. We first develop a prediction model using machine learning techniques: general linear model, Multivariate Adaptive Regression Splines (MARS), artificial neural networks, classification and regression tree. The best model is built by the classification and regression tree with an accuracy of 84%. Based on the predicted values, we redefine the appointment grids, which are evaluated using the sequence rules and the patient management rules. The best solutions are selected based on the number of treated patients, the waiting time and the overtime. By applying the new approach, we will have a 30% increase in patients per day.

This thesis addresses the problem of scheduling patient appointments and technologist schedules. In most health centers, these two elements are treated separately. In our study, we develop two versions of an optimization model: sequential and integrated. The first version consists of developing the schedules of the technologists initially and designing the schedules of appointments thereafter; however, the second version combines the two elements. We aim to ensure optimal staff allocation while maximizing the use of machines, the number of

patients treated, and the stability of schedules and grids. Our approach is evaluated based on real data from the MRI department of the CHUM, by modifying the technologists working rules and the planning construction methodes. Although the sequential version is the simplest, it gives good results according to certain indicators (such as the stability of the technologist's schedules), to the detriment of other indicators (such as the stability of the appointment schedule), unlike the integrated method that considers all factors simultaneously. By conducting an analysis that clarifies the impact of each solution versus the complexity of the implementation, we guide administrators to the best solutions and the optimal plan for their application.

The last contribution of this thesis concerns the rescheduling of patients in radiotherapy under machine breakdown. We develop an optimization model and a heuristic. They are applied within a well-defined framework comprising all the sequences of priority for rescheduling decisions, which are: delaying a patient, overbooking a patient, using the overtime to schedule a patient. In our case study, we use historical data from the CICL concerning the treatment times, the number of treatment sessions, and the due dates of treatment plans. Given the lack of documentation related to machine breakdowns, the proposed approach is evaluated using the scenarios generated by modifying the duration of the breakdown and the schedule filling rate. The solutions are compared by: the number of delayed patients, the waiting time, and the overtime. We present a detailed analysis of the results that allows managers to understand the impact of the decisions made on performance.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ABRÉVIATIONS

CHAID	Chi-squared Automatic Interaction Detection
CHUM	Centre hospitalier de l'Université de Montréal
CICL	Centre intégré de cancérologie de Laval
CR tree	Arbre de classification et de régression (*classification and regression*)
CRISP-DM process	CRoss Industry Standard Process for Data Mining
CT	Computed tomography
GRASP	Greedy Random Adaptive Search Procedure
IRM	Imagerie par résonance magnétique
Linac	Linear accelerator
MARS	Multivariate adaptive regression splines
MIP	Mixed-integer linear program
MRI	Magnetic resonance imaging
OCDE	Organisation de coopération et de développement économiques
PIB	Produit intérieur brut
SMF	Smallest-mean-first
SVF	Smallest-variance-first
TDM	Tomodensitométrie

CHAPITRE 1 INTRODUCTION

La croissance de la population pousse les gouvernements à faire des investissements importants dans le domaine de santé afin d'assurer les soins nécessaires aux patients. Selon l'Organisation de coopération et de développement économiques (OCDE), le Canada dépense 6 666 $ par personne en 2019 pour la santé, ce qui correspond à 10,8 % du produit intérieur brut (PIB). Les États-Unis sont un des pays de l'OCDE qui dépense le plus en santé, à 13 590 $ par personne, 16,8 % du PIB. En outre, le vieillissement de la population participe à l'augmentation de ces dépenses. En effet, 45 % des dépenses publiques de santé des provinces et territoires au Canada sont dédiées aux personnes de 65 ans et plus qui ne représentent qu'environ 18 % de la population canadienne (Institut canadien d'information sur la santé, 2021a). Au Québec, les dépenses en services de santé et services sociaux représentent plus du tiers des dépenses publiques, ils valent 32,7 G$ en 2015-2016 et 38,4 G$ en 2016-2017 avec une augmentation de 1,95 % (Hébert *et al.*, 2017). Une grande partie des dépenses est consacrée aux ressources humaines. En effet, près des deux tiers du financement en santé et services sociaux au Québec sont dédiés aux travailleuses et travailleurs de la santé. L'effectif du personnel atteint 300 000 en 2015-2016, dont 200 000 professionnelles salariées, telles que les infirmières, les préposées aux bénéficiaires et autres techniciennes ou professionnelles de la santé (Hébert *et al.*, 2017).

Le coût élevé de certains traitements empêchera les gouvernements de faire davantage d'investissements, ainsi que le nombre limité des ressources humaines ne permet pas l'exploitation maximale des équipements médicaux. Par conséquent, il n'est pas évident de garantir les délais recommandés pour tous les patients, la situation est très critique quand il s'agit des maladies mortelles telles que le cancer. En effet, le cancer est la cause principale de décès au Canada. 233 900 nouveaux cas de cancer et 85 100 décès sont estimés en 2022 selon le rapport annuel de la Société canadienne du cancer. Il est prévu que deux Canadiens sur cinq seront atteints de cancer au cours de leur vie, et un Canadien sur quatre mourra du cancer (Société canadienne du cancer, 2021).

Les patients dans les cliniques ambulatoires ont besoin d'accéder aux soins dans des délais spécifiques afin d'assurer le bon traitement au bon moment. En 2004, les Premiers ministres canadiens se sont engagés à réduire les temps d'attente dans quelques domaines prioritaires, parmi lesquels nous citons le traitement du cancer et l'imagerie diagnostique (Institut canadien d'information sur la santé, 2022). Le temps d'attente pour une radiothérapie s'est amélioré : 11 jours ou moins pour 50 % des patients, et 21 jours pour 90 % des patients, sachant que le délai maximum recommandé est 28 jours (Institut canadien d'information sur la santé, 2021b). Pour l'imagerie médicale, les temps d'attente recommandés pour les patients non urgents sont largement dépassés. Les Canadiens pourraient attendre en moyenne 4,8 semaines pour la tomodensitométrie (TDM) et 9,3 semaines pour l'imagerie par résonance magnétique (IRM) (Bacchus et Mackenzie, 2019). Le nombre d'appareils de TDM et d'IRM au Canada est inférieur à la médiane des unités par habitant de l'OCDE. Afin d'assurer le service aux patients dans des délais acceptables, l'Association canadienne des radiologistes recommande de consacrer 1,5 milliard de dollars sur 5 ans au budget fédéral 2022 pour renouveler les équipements d'imagerie diagnostique et interventionnelle (Gilles *et al.*, 2022), sachant que ces coûts de renouvellement sont estimés à 4,4 milliards de dollars d'ici 2040 (Association canadienne des radiologistes, 2020). Bien que les équipements soient restreints, la disponibilité de plus de technologues permettra de bien exploiter ces machines en élargissant leur temps d'utilisation. Or, les radiologistes canadiens confirment que les ressources humaines en santé représentent la principale limite à l'augmentation de la capacité d'imagerie (Gilles *et al.*, 2022).

Le déséquilibre entre la demande d'accès aux soins et la capacité des centres médicaux devient de plus en plus important surtout avec la croissance continue de la population, ce qui influe sur la satisfaction des patients et les dépenses médicales telles que le temps supplémentaire travaillé par le personnel. Les gestionnaires des cliniques ambulatoires font face aux longues listes d'attente des patients et un budget médical limité. Donc, ils sont obligés de faire une gestion efficace des ressources en réduisant les coûts et en assurant une bonne qualité de service.

La présente thèse consiste à concevoir de nouveaux outils d'aide à la décision, permettant d'optimiser la gestion des ressources, et aussi d'améliorer le processus de prise des rendez-vous, en tirant profit des méthodes d'apprentissage machine et des outils de la recherche opérationnelle. Dans cette thèse, nous allons proposer une modélisation d'un système de prise de rendez-vous en prenant en compte des décisions qui s'intègrent dans les deux niveaux : tactique et opérationnel.

Afin de bien illustrer l'intérêt de la méthodologie proposée et son impact sur l'amélioration de la gestion des rendez-vous de patients et des horaires des technologues, le projet sera appliqué sur deux types de centres de services ambulatoires : le département de la radiothérapie du CICL, et le département d'imagerie médicale du CHUM. En effet, le processus de la radiothérapie est complexe. Il est constitué de deux phases principales : le prétraitement et le traitement. La première phase comporte quatre étapes : le moulage, l'imagerie par la TDM, la dosimétrie, et la simulation. Elle est caractérisée par l'utilisation de différentes ressources à la fois, la recirculation des patients (un patient a besoin d'une ressource plus d'une fois) et la non-disponibilité des ressources d'une façon continue tout au long de l'horizon de planification (Castro et Petrovic, 2012). En revanche, le processus d'imagerie médicale est plus simple, nous pouvons le voir comme une activité dans la trajectoire de la radiothérapie (par exemple, l'imagerie par la TDM).

Deux types de stratégie de planification des rendez-vous de patients sont utilisés : le système avec bloc et le système sans bloc. Le premier partitionne la journée en un nombre fixe de blocs de la même durée, tandis que le deuxième comporte des intervalles de rendez-vous de durées différentes. Plusieurs centres externes, tels que le CICL, adoptent la stratégie de gestion des rendez-vous avec blocs vu sa facilité. Ils supposent que les patients ont les mêmes caractéristiques, donc, ils peuvent leur attribuer des créneaux horaires de même largeur estimée en se basant sur la durée de traitement moyenne. En fait, les patients sont différents en termes de la catégorie de la maladie, la nature de la visite médicale, la mobilité, etc. Le fait de considérer que les patients ont la même durée de service peut bloquer le potentiel d'amélioration du système de prise des rendez-vous, et engendrer des coûts liés à la sous-utilisation des ressources, le temps supplémentaire, etc.

Dans la plupart des centres d'imagerie médicale, tels que le département d'imagerie du CHUM, les créneaux horaires attribués aux patients sont variables en fonction de type de l'examen, la machine utilisée, la partie du corps traité, etc. ; par contre, les grilles des rendez-vous sont standardisées et les heures de fonctionnement de chaque machine sont prédéfinies. En outre, les plannings des technologues sont sélectionnés dépendamment de la salle affectée. Le nombre de technologues alloués à chaque machine est fixe, il dépend du quart de travail et non pas de la catégorie des examens. La réalisation de quelques examens demande la présence du personnel médical spécifique, ou bien elle nécessite plus d'un personnel soignant. En effet, l'élaboration intégrée des grilles des patients et des horaires des technologues représente la modélisation la plus réaliste et elle peut ouvrir de nouvelles pistes d'amélioration de la performance, mais, elle amplifiera la complexité du problème. Pour simplifier la réalisation de ces deux tâches, deux agents différents élaborent mensuellement les grilles des rendez-vous et les horaires des technologues d'une manière séparée. Les modifications produites au niveau des machines, des examens, ou des technologues nécessitent la communication entre les deux agents et des aller-retour, pour faire les ajustements nécessaires tout en réalisant le minimum des changements possible.

La gestion de l'incertitude est une autre composante importante dans le système de santé. L'occurrence des événements imprévus tels que les pannes des machines, les patients urgents, les annulations de traitement, etc. impacte les plannings élaborés initialement, ce qui rend leur replanification une opération nécessaire à considérer. Or, la planification des traitements en radiothérapie confronte plusieurs défis. En effet, il faut respecter le plan de traitement de chaque patient. De plus, les centres de radiothérapie veillent à planifier les séances de traitement quotidiennement, à la même heure et la même machine. En outre, ils font de leur mieux pour assurer les préférences des patients. Donc, en cas d'apparition d'événements imprévus, l'exécution des nouveaux plannings optimaux est une tâche complexe. Pour ce faire, les administrateurs cherchent des solutions rapides sans évaluer leur efficience, ils replanifient les séances annulées sans suivre une procédure bien définie.

L'objectif principal de cette thèse est de proposer des outils d'aide à la décision qui permettront d'analyser la situation actuelle et d'améliorer le système de planification des rendez-vous et des horaires des technologues dans les centres de santé. Pour ce faire, nous utilisant des techniques d'apprentissage machine et de la recherche opérationnelle. Le contenu de cette thèse est divisé en trois objectifs.

- Le premier objectif consiste à développer un modèle de prédiction du temps de service des patients en radiothérapie en utilisant des techniques d'apprentissage machine. En nous basant sur les durées prédites, nous élaborons de nouvelles grilles des rendez-vous des patients qui sont évaluées à l'aide des règles de gestion et de séquence des patients.

- Le deuxième objectif consiste à comparer deux versions d'un modèle d'optimisation pour la confection des grilles des patients et des horaires des technologues en imagerie médicale : intégré et séquentiel. Les deux versions du modèle sont évaluées en modifiant les règles de travail des technologues et les méthodes de construction des plannings. En outre, nous visons à étudier l'impact du nombre des technologues alloués aux examens sur le nombre total des patients traités.

- Le troisième objectif consiste à proposer des stratégies pour replanifier les patients en radiothérapie suite à une panne de machine. Pour ce faire, un modèle d'optimisation et une heuristique sont développés et évalués en changeant les priorités des décisions de replanification.

La thèse comporte huit chapitres. Le chapitre 2 concerne une revue de la littérature existence. Le chapitre 3 présente la méthodologie appliquée dans cette thèse. Les chapitres 4, 5, et 6 présentent respectivement les articles scientifiques liés au premier, deuxième et troisième objectif. Le premier article «Patient scheduling based on a service-time prediction model : A data-driven study for a radiotherapy center», et le deuxième article «On integrating patients appointment grids and technologist schedules in a radiology center» sont publiés dans Health Care Management Science, le troisième article «Radiotherapy patient rescheduling under machine breakdown» est soumis au même journal. Le chapitre 7 propose une discussion générale qui contient une synthèse des travaux que nous avons réalisés, et quelques perspectives de recherches futures. Nous achevons cette thèse par une conclusion dans le chapitre 8.

CHAPITRE 2 REVUE DE LITTÉRATURE

Dans ce chapitre, nous présentons une revue sur les travaux de recherche liés à la planification des patients et du personnel en santé. La première section traite l'exploitation des techniques d'apprentissage machine en santé pour l'extraction de l'information utile. La deuxième section propose les méthodes de la planification des patients dans les cliniques ambulatoires, notamment, en radiothérapie et imagerie médicale. La troisième section concerne la replanification en santé. La quatrième section présente la confection des horaires du personnel en santé. La dernière section regroupe les travaux qui intègrent deux éléments de planification en santé.

2.1 L'application des techniques d'apprentissage machine en santé

Le stockage de données médicales permet de construire un historique de données riche, qui motive plusieurs chercheurs à mener des études axées sur les données, et appliquer des méthodes avancées d'exploration de données afin d'extraire de l'information et de la connaissance. Les techniques d'apprentissage machine ont été largement utilisés dans le milieu de la santé pour la prédiction d'une maladie (Jain et Singh, 2016), l'évaluation de l'efficacité du traitement, et le regroupement des patients à risque élevé (Koh *et al.*, 2011). Dans notre étude, nous sommes intéressés par les travaux qui servent à améliorer la planification des patients dans les hôpitaux et les services de soins ambulatoires.

De nombreuses études sont concentrées sur la prédiction du « No-show » des patients pour réduire les coûts associés dans les centres de santé. Les modèles prédictifs du « No-show » sont généralement utilisés dans la surréservation des patients ambulatoires. Harris *et al.* (2016) ont développé un modèle d'analyse prédictive du « No-show » qui combine un modèle de régression et l'approximation fonctionnelle en utilisant la somme des fonctions exponentielles. D'autres techniques de prédiction telles que la régression logistique et l'arbre de classification ont été utilisées pour évaluer le modèle proposé. Moradi *et al.* (2022) ont présenté une étude axée sur les données pour la planification des rendez-vous des patients qui tient compte de leurs comportements. Leur approche est composée de trois étapes principales. Ils ont commencé par utiliser des techniques d'apprentissage machine pour classer les patients en cas du « No-show » et pour prédire leur retard. Dans un deuxième temps, ils ont priorisé les patients en fonction de la gravité de leur maladie et de la prolongation du traitement. Ces prédictions et priorités ont été prises en compte dans un modèle de programmation linéaire mixte en nombres entiers pour la planification des patients en minimisant le temps d'inactivité, les heures supplémentaires et le temps d'attente des patients. Ils ont réussi à réduire 30 % des coûts totaux du centre. Topuz *et al.* (2018) ont prédit la probabilité du « No-show » en pédiatrie par 18 variables de prédiction. Srinivas et Ravindran (2018) ont utilisé des classificateurs d'apprentissage machine (Régression logistique, Réseaux de neurones, Forêts aléatoires, méthodes d'ensemble (Gradient Boosting, Stacking)) pour prédire le risque du « No-show » d'un patient. Sur la base des résultats obtenus, ils ont proposé des règles de planification en intégrant des règles d'enchaînement et des règles de surréservation. À l'aide d'une approche d'apprentissage machine basée sur les données, Srinivas et Salah (2021) ont développé un modèle de classification pour prédire le « No-show », et un algorithme de régression pour estimer les durées de consultation pour une clinique de cardiologie. Ils ont évalué la valeur ajoutée de l'intégration de ces prédictions pour la prise de rendez-vous. Simsek *et al.* (2020) ont prédit le « No-show » des patients en utilisant un réseau de neurones artificiel, et en sélectionnant des variables à l'aide d'un algorithme génétique et d'un recuit simulé. Dantas *et al.* (2019) ont développé des modèles prédictifs pour le « No-show » des patients dans les cliniques de chirurgie bariatrique, et ils ont identifié les facteurs démographiques et les informations de planification les plus importants. Fan *et al.* (2021) ont élaboré des modèles de prédiction du « No-show » des patients pour les rendez-vous ambulatoires en ligne à l'aide de plusieurs algorithmes d'apprentissage machine : régression logistique, k-plus proche voisin, arbre de décision, forêt aléatoire et les méthodes d'ensemble (bagging, boosting).

Erekat *et al.* (2020) ont utilisé des algorithmes d'exploration de données pour prédire les annulations de chirurgies afin de réduire le nombre de créneaux annulés. Sur la base des résultats de la prédiction, ils proposent des stratégies de planification des chirurgies, qui sont évaluées par rapport au calendrier de l'état actuel à l'aide d'un modèle de simulation à événements discrets. Ils confirment l'efficacité de leur approche de planification pour augmenter l'utilisation des salles d'opération.

Pour la prédiction de la réadmission des patients à l'hôpital, la plupart des chercheurs ont appliqué un seul modèle de prédiction à l'aide de la régression logistique, contrairement à Golmohammadi et Radnia (2016) qui ont appliqué les réseaux de neurones, l'arbre de classification et de régression et la méthode CHAID (Chi-squared Automatic Interaction Detection). Tous les modèles construits fonctionnent bien avec une précision globale supérieure à 80 %. De même, Braga *et al.* (2014) ont proposé plusieurs modèles pour la prédiction de la réadmission, y compris la machine à vecteurs de support, le classifieur bayésien naïf, et les arbres de décision.

Alkhatib et Alahmar (2021) ont présenté une revue de littérature sur la prédiction de la durée de séjour à l'hôpital à l'aide l'apprentissage machine et des approches statistiques, en se concentrant sur les patients d'AVC. Pereira *et al.* (2016) ont comparé plusieurs modèles prédictifs du temps d'attente entre le triage et l'admission médicale pour les soins de maternité. Les modèles sont générés par différentes techniques d'apprentissage machine : les arbres de décision, la classification bayésienne naïve, la machine à vecteurs de support, les modèles linéaires généralisés et les réseaux de neurones. Ils ont déduit que le meilleur modèle a été obtenu en appliquant les arbres de décision. Kuo *et al.* (2020) ont présenté une prédiction précise et personnalisée des temps d'attente dans les services d'urgence en utilisant des algorithmes d'apprentissage machine.

Afin de prédire le temps de service avec précision, Golmohammadi (2021) ont utilisé les réseaux de neurones en supposant l'hétérogénéité des caractéristiques des patients. La solution proposée a prouvé son efficacité vis-à-vis d'un modèle d'ordonnancement commun. Huang et Marcak (2013) ont appliqué un arbre de décision pour classer les patients en fonction de leurs caractéristiques. Les résultats ont été exploités dans l'attribution des intervalles de temps appropriés. Cette approche augmente l'utilisation des radiographes et l'accès des patients.

Nous pouvons conclure que la plupart des modèles de prédiction élaborés sont dédiés pour la prédiction des variables de sorties binaires telles que le « No-show », cependant, il existe quelques travaux qui s'intéressent à la prédiction du temps de service ou du temps d'attente dans des applications autres que la radiothérapie.

2.2 La planification des rendez-vous de patients dans les centres de soins ambulatoires

Les systèmes de gestion de rendez-vous de patients ambulatoires sont parmi les sujets qui ont attiré l'attention de plusieurs chercheurs dans les dernières années. Trois revues de littérature sont considérées importantes pour avoir un aperçu global sur le sujet, qui sont : Cayirli et Veral (2003), Gupta et Denton (2008) et Ahmadi-Javid *et al.* (2017). La première revue a été focalisée sur les approches de modélisation et formulation du problème de planification des patients ambulatoires. Elles ont été réparties en trois groupes : programmation mathématique, simulation et théorie des files d'attente. Gupta et Denton (2008) ont décrit les types de systèmes de soins de santé, en mettant l'accent sur les facteurs qui compliquent la planification des rendez-vous. Plus récemment, Ahmadi-Javid *et al.* (2017) ont fait une mise à jour à ces deux revues de littérature. Ils ont classifié les décisions prises pour modéliser un système de prise de rendez-vous en trois : stratégiques, tactiques, et opérationnelles. Les travaux qui s'intéressent aux décisions stratégiques sont rares, elles sont prises généralement comme des éléments d'entrée. Le niveau stratégique comporte la détermination de la politique d'accès aux soins (accès libre ou traditionnel), le nombre de ressources, le type de planification (online ou offline), etc. Par contre, le deuxième niveau consiste à préciser la longueur des intervalles des rendez-vous, leur nombre par session, etc. Les décisions opérationnelles sont les plus étudiées par les auteurs, elles contiennent l'affectation des patients aux ressources, la sélection du jour ou l'heure du rendez-vous, la séquence des patients, etc. En outre, une étude intéressante a été réalisée par Cayirli *et al.* (2006) qui confirme l'impact de l'intégration de la classification des patients sur l'amélioration des systèmes de prise des rendez-vous. Ils ont proposé six règles de séquence des patients en distinguant les patients en nouveau/retourné. Ces règles ont été simulées en les combinant avec sept règles de planification des rendez-vous, en tenant en compte des facteurs d'environnement, tels que le nombre de patients par session, la probabilité du « No-Show », le coefficient de variation des temps de service et la ponctualité des patients. Ils ont prouvé l'importance de la considération des règles de séquence des patients dans les systèmes de gestion des rendez-vous dans les cliniques ambulatoires par rapport aux règles de planification des rendez-vous.

Dans la présente revue de littérature, nous mettons l'accent sur les travaux connexes à la planification des rendez-vous des patients ambulatoires aux centres de radiothérapie, ainsi qu'aux centres d'imagerie médicale.

2.2.1 Les centres de radiothérapie

Le nombre des articles publiés sur la gestion des opérations de soins contre le cancer a augmenté rapidement ces dernières années. Hadid *et al.* (2022) ont évalué la structure des connaissances et le développement de la recherche dans ce domaine. Différentes approches ont été adoptées pour la modélisation du problème de planification des rendez-vous de patients dans les centres de radiothérapie. Vieira *et al.* (2016) ont confirmé que la programmation mathématique est la plus utilisée dans ce sujet de recherche ; par contre, il existe d'autres chercheurs qui préfèrent l'application des heuristiques et des méthaheuristiques. Les objectifs considérés dans la littérature visent à maximiser le nombre de patients traités, minimiser le temps d'attente des patients, ou bien minimiser le retard par rapport à la date d'échéance.

Conforti *et al.* (2008) ont développé un modèle mathématique qui vise une planification optimale des traitements en radiothérapie. Son objectif est de maximiser le nombre de patients traités en prenant en considération leurs priorités. Ils ont intégré dans un modèle basic toutes les contraintes réelles d'un problème de la planification en radiothérapie. Par la suite, ils ont conçu un autre modèle en ajoutant la possibilité de changer les rendez-vous de patients qui ont été déjà planifiés. L'application de ce modèle sur un cas réel a donné de bons résultats : le nombre de patients traités a été augmenté, et la durée entre le diagnostic et le début de traitement a été minimisée. Conforti *et al.* (2011) ont étendu ce travail en tenant compte d'autres exigences telles que la disponibilité des patients. En outre, ils ont donné plus de flexibilité au modèle par la relaxation de la contrainte d'avoir toutes les séances de traitement dans des jours consécutifs. L'objectif du modèle amélioré est non seulement la maximisation du nombre de patients traités, mais aussi la minimisation du retard par rapport à la date d'échéance du début de traitement. Conforti *et al.* (2010) ont proposé un programme d'optimisation linéaire qui maximise le nombre de patients planifiés, et favorise l'efficience des accélérateurs linéaires. Ces derniers sont considérés égaux à deux, contrairement à Conforti *et al.* (2008) et Conforti *et al.* (2011) qu'ils ont fait l'hypothèse de l'utilisation d'un seul accélérateur linéaire.

Vieira *et al.* (2020) a proposé un modèle de programmation linéaire mixte en nombres entiers pour la planification des rendez-cous des patients en radiothérapie en prenant compte des préférences de fenêtre de temps. (Boonmee *et al.*, 2021) ont présenté un modèle qui considère différents sites de traitement, différentes technologies de radiothérapie, différents types de patients, le temps disponible des médecins, la planification de la chimiothérapie, et le processus de simulation.

Les différents modèles proposés sont généralement statiques, et par conséquent, ils fournissent une solution myope, qui ne tient pas compte de la variabilité, l'incertitude, et la dynamique de l'environnement. Saure *et al.* (2012) ont développé une politique dynamique basée sur le processus de décision de Markov. Leur objectif est l'affectation des rendez-vous multiples aux patients en considérant leurs priorités, tout en minimisant le coût total associé à la réservation d'un rendez-vous. Les résultats montrent que la solution proposée performe mieux que la politique myope. Le pourcentage des nouveaux patients traités dans dix jours a été augmenté de 73 % à 96 %. En se basant sur le même modèle de Saure *et al.* (2012), Gocgun (2018) a considéré la possibilité d'annulation des rendez-vous des patients. Son travail confirme que la solution proposée est la meilleure à utiliser quand la capacité du système étudié est restreinte.

L'incertitude est présente dans le processus de la radiothérapie, elle est liée à l'arrivée des patients, la durée de traitement, la panne des machines, etc. La prise en compte de cette composante permet que les solutions proposées soient bien adéquates pour leur application en réalité (Legrain, 2015). Legrain *et al.* (2015) ont développé une méthode hybride qui combine l'optimisation stochastique et l'optimisation « online » , dans le but de déterminer une bonne politique de planification des patients au CICL. Les algorithmes « online » permettent de prendre des décisions rapides en temps réel tout en considérant la bonne qualité de la solution offerte. Les auteur ont présenté un algorithmes « offline » qui performe mieux que celui du centre, par contre, il ne prend pas en considération l'incertitude de l'arrivée des patients et leurs priorités. Par la suite, ils ont proposé un algorithme stochastique « online » , qui utilise des scénarios qui tiennent compte des événements imprévus. Braune *et al.* (2021)

ont présenté un modèle d'optimisation stochastique pour une planification des rendez-vous en radiothérapie, en considérant l'incertitude dans les durées d'activité. Ils ont résolu le problème à l'aide d'une heuristique en combinant un algorithme génétique et une simulation de Monte-Carlo.

Vu leur simplicité et également leur efficacité, les heuristiques ont été utilisées dans plusieurs travaux. Petrovic *et al.* (2006) ont présenté deux algorithmes d'affectation des rendez-vous de patients déjà priorisés selon différentes catégories. Le premier planifie les patients vers l'avant en commençant par la première date possible de début de traitement. Par contre, le deuxième les programme vers l'arrière à partir de la dernière date possible de début de traitement. Les algorithmes proposés ont été évalués en se basant sur les critères suivants : le nombre de patients dont la durée d'attente cible n'est pas respectée, la durée totale des délais d'attente des patients, et le nombre d'interruptions des traitements programmés. Il a été recommandé dans cet article d'essayer la combinaison entre les deux algorithmes.

D'autres chercheurs ont choisi l'application des méthodes plus complexes que les heuristiques simples, qui sont les métaheuristiques. Petrovic *et al.* (2009) ont implémenté un algorithme génétique pour la planification des patients en radiothérapie, en définissant deux objectifs qui minimisent le temps d'attente moyen des patients, et le retard moyen par rapport à la date d'échéance du premier traitement. Le même algorithme a été utilisé par Petrovic et Castro (2011) pour la planification dans la phase de prétraitement. La performance a été évaluée par le temps d'attente cible du patient et le temps d'inactivité des ressources. Kapamara *et al.* (2006) traitent un problème réel de planification des patients en radiothérapie, qui peut être considéré un problème «job shop» stochastique dynamique NP difficile. Pour cette raison, ils ont proposé un ensemble d'approches de résolution de ce type de problèmes. Ils ont commencé par l'application des règles d'affectation telles que le choix du patient avec le temps de traitement le plus court ou bien le temps de traitement le plus long, etc. Par la suite, ils ont utilisé des métaheuristiques, notamment, le recuit simulé, GRASP (Greedy Random Adaptive Search Procedure), l'algorithme génétique et la recherche taboue. Ils ont prouvé que cette dernière performe mieux que les autres.

D'autres études sont basées sur l'hybridation en combinant les heuristiques avec les métaheuristiques, ou bien avec la programmation mathématique. Dans l'article de Petrovic et Leite-Rocha (2008), quatre approches constructives ont été développées pour la planification du traitement de la radiothérapie. L'objectif global est la minimisation du retard moyen pondéré des patients (le retard est calculé par rapport au temps d'attente cible). De plus, ils ont présenté un algorithme basé sur GRASP pour améliorer les solutions créées par les approches proposées. Les résultats montrent que l'utilisation de GRASP n'est pas justifiée. Dans le même contexte, Castro et Petrovic (2012) ont modélisé la planification des rendez-vous des patients dans la phase de prétraitement en radiothérapie sous forme d'un problème à trois objectifs, en combinant les règles d'affectation et la programmation mathématiques. En effet, les trois objectifs sont hiérarchiques, et la résolution du modèle en CPLEX était longue. De ce fait, les règles de planification ont été utilisées pour construire une solution initiale de la première phase.

À la fin de cette sous-section, nous mentionnons que deux stratégies de planification peuvent être appliquées en radiothérapie : avec bloc et sans bloc. Conforti *et al.* (2010) ont défini la planification avec bloc par celle qui attribuent les créneaux horaires de même durée à tous les patients, contrairement à la planification sans bloc qui permet de donner des rendez-vous de durée variable selon le cas du patient traité. Bien que cette dernière soit plus efficiente, plusieurs centres de radiothérapie adoptent la stratégie avec bloc dans la planification des rendez-vous de patients. Conforti *et al.* (2010) ont confirmé que l'assignation des blocs de rendez-vous uniformes ne représente pas bien la charge de travail réelle. Nous déduisons que la possibilité d'amélioration de la performance augmente en utilisant le surplus de temps assigné aux patients.

2.2.2 Les centres d'imagerie médicale

La majorité des articles qui traitent la planification des patients dans les centres d'imagerie médicale utilisent des méthodes de la modélisation, la simulation, et l'optimisation. Van Sambeek *et al.* (2011) ont modélisé le processus actuel, et ils ont appliqué la simulation à événements discrets afin de minimiser le temps d'accès à l'imagerie par résonance magnétique. Les patients ont été catégorisés en 15 groupes selon le test demandé, cette étude a proposé la réduction du nombre de catégories à quatre. Walter (1973) a construit un modèle de simulation par ordinateur. Il a évalué l'impact des changements appliqués sur des paramètres dans un département de radiologie (tels que le nombre de patients avec et sans rendez-vous) sur la performance. Cette dernière est mesurée par le temps moyen d'attente du patient et le temps d'inactivité du médecin. Il a montré que l'efficience s'améliore quand le nombre de patients avec rendez-vous augmente. En outre, il a prouvé que le temps de service du patient dépend de plusieurs variables comme son âge, son origine (ambulatoire ou non) et sa mobilité. Elkin *et al.* (2012) ont réalisé un sondage en interrogeant par téléphone toutes les installations de mammographie certifiées par le gouvernement fédéral et situées dans six villes aux États-Unis, dans le but d'évaluer l'impact de la capacité (nombre de machines de mammographie) sur le temps d'attente pour avoir un rendez-vous. La relation entre ces deux variables a été modélisée par la régression logistique multinomiale, en catégorisant le temps d'attente. Les résultats montrent que dans les zones à faible capacité, la demande de rendez-vous dépasse l'offre d'une grande marge, ce qui crée de longs temps d'attente.

D'autres chercheurs ont utilisé des méthodes plus complexes pour la gestion des patients. Kolisch et Sickinger (2008) ont traité un problème d'allocation dynamique de deux scanners parallèles dans un département de radiologie à une demande stochastique de différents groupes de patients. Le problème a été modélisé par un processus de décision de Markov. Dans le domaine médical, les règles de décision qui peuvent être appliquées à la main sont préférables et faciles à utiliser. Par conséquent, trois règles de décision communes dans les hôpitaux ont été évaluées sous trois calendriers de rendez-vous avec trois différents scénarios. Ils ont recommandé l'utilisation de la règle «premier arrivé premier servi», puisqu'elle est facile à implémenter, et elle performe mieux que les autres dans la plupart des cas. En outre, Patrick et Puterman (2007) ont développé une approche qui maximise l'utilisation et réduit les temps d'attente pour une ressource de diagnostic, en présence d'une demande incertaine, avec plusieurs niveaux de priorité : les patients hospitalisés ont une haute priorité, et les patients ambulatoires ont une faible priorité. La méthode proposée introduit une troisième classe qui est les patients hospitalisés avec une faible priorité, qui ont un jour de flexibilité dans la planification de leurs rendez-vous. Le modèle d'optimisation minimise le nombre prévu de créneaux vides, tout en respectant une limite de temps supplémentaire. Les résultats de la simulation prouvent que le fait de considérer 10 % des patients hospitalisés à faible priorité donne une réduction significative du temps d'attente des patients ambulatoires. Dans un travail postérieur, Patrick *et al.* (2008) ont présenté une méthode de planification dynamique des patients à différentes priorités, pour accéder à la ressource d'imagerie sans dépasser les temps d'attente cible. Il s'agit d'une modélisation par le processus de décision de Markov, en résolvant le programme linéaire équivalent par la programmation linéaire approximée.

Chen *et al.* (2022) ont utilisé l'optimisation et la simulation pour trouver la durée appropriée de l'intervalle de temps pour les patients subissant un examen d'échographie, en tenant compte de plusieurs types de patients, plusieurs salles d'examen, plusieurs parties du corps, et de plusieurs technologues en radiologie. La solution obtenue permet une meilleure affectation des patients aux salles d'examen, équilibre la charge de travail des technologues en radiologie, assure des taux d'utilisation élevés des équipements, et réduit les temps d'attente des patients.

Choy *et al.* (2018) ont confirmé que l'application des techniques d'apprentissage machine peut optimiser la planification des patients en radiologie, en identifiant par exemple les patients avec une grande probabilité de la non-présentation au rendez-vous. Glover IV *et al.* (2017) ont identifié les patients ayant une forte probabilité de ne pas se présenter à leur rendez-vous de radiologie.

Les travaux cités concernent la planification des patients en imagerie médicale sans prendre en parallèle d'autres composantes du processus telles que la planification du personnel. En effet, la considération de ces deux éléments de planification à la fois peut augmenter le potentiel de l'amélioration de la performance.

2.3 La planification du personnel

La planification du personnel a été largement étudiée dernièrement. Nous considérons trois revues de littérature liées à ce sujet : Van den Bergh *et al.* (2013); Ernst *et al.* (2004); Alfares (2004). Ils ont présenté la classification du problème en littérature, les méthodes de résolution, et les domaines d'application. La planification du personnel est définie par le processus de construction des horaires de travail du personnel d'une organisation pour satisfaire sa demande. Ernst *et al.* (2004) ont proposé un cadre général du problème en le décomposant en un processus de six modules : modélisation de la demande, planification des jours de congé, planification des quarts, construction de la ligne de travail, affectation des tâches, et affectation du personnel. Baker (1976) ont classifié le problème en : planification des quarts de travail, planification des jours de congé, et planification des tournées qui regroupe les deux premiers types de planification. Les méthodes de résolution du problème peuvent être placés sous différentes catégories : la programmation mathématique, la programmation par contraintes, la simulation, les heuristiques et les métaheuristiques, etc.

La planification du personnel est primordiale dans plusieurs domaines, nous sommes intéressés par son application en santé. En effet, la construction des horaires des infirmiers et des médecins a attiré l'attention de plusieurs chercheurs (Burke *et al.*, 2004; Erhard *et al.*, 2018). Afin de capturer la complexité du problème en réalité, plusieurs aspects doivent être considérés.

L'élaboration des horaires du personnel en santé comporte l'intégration des contraintes de plusieurs types : dures et souples. Il s'agit de respecter les contraintes de couverture, de temps et de réglementation, ainsi que l'équité et les préférences des employés. La modélisation du problème confronte de nombreux défis vu l'hétérogénéité des employés qui ont différents compétences et contrats de travail. De plus, il existe deux types de quarts de travail : prédéfini et flexible. Brunner *et al.* (2009) ont présenté le problème de planification des médecins en générant des lignes de travail individuel pour chacun d'eux avec des quarts de travail flexibles. Ils ont considéré des temps de début des quarts flexibles avec des longueurs variables, des pauses, du temps supplémentaire, et des services sur appel. Afin de trouver des solutions de bonne qualité, ils ont utilisé l'algorithme Brunch-and-Price qui permet de résoudre de grandes instances représentant la taille réelle du problème (Brunner *et al.*, 2010). Stolletz et Brunner (2012) ont proposé un modèle de construction des horaires des médecins avec des quarts de travail flexibles en introduisant l'équité.

Il y a une grande similarité entre la planification des horaires des infirmiers et la planification des horaires des technologues. Sauf que dans la phase de traitement en radiothérapie et en imagerie médicale, les technologues représentent la ressource personnelle principale et la plus utilisée. Il n'existe que quelques papiers dans la littérature qui étudient la planification des horaires des technologues. Chen *et al.* (2016) ont traité un problème intégré d'allocation et de planification du personnel médical. Ils ont développé un algorithme en deux étapes qui détermine le nombre minimal du personnel dans la première étape, et construit les horaires dans la deuxième étape. Ils ont appliqué leur approche dans une étude de cas qui concerne les technologues en radiologie. Yuura *et al.* (2017) ont présenté un modèle de programmation en nombres entiers pour la planification des technologues en radiographie en considérant leurs compétences et formations. Vieira *et al.* (2018) ont optimisé l'affectation des technologues en radiothérapie à plusieurs opérations en tenant compte des arrivées stochastiques des patients.

La planification des rendez-vous du personnel en santé est une tâche complexe, notamment avec l'intégration de tous les aspects réels du problème. La considération d'autres éléments de planification peut augmenter la complexité de la représentation et la résolution du problème. Pour ce faire, la plupart des chercheurs traitent ce sujet d'une manière indépendante.

2.4 La planification intégrée en santé

Dans la littérature, il n'existe que quelques travaux qui intègrent deux éléments de la planification en santé. Ogulata *et al.* (2008) ont développé un modèle mathématique hiérarchique pour la planification hebdomadaire des physiothérapeutes. La première étape détermine le nombre maximal des patients sélectionnés pour la physiothérapie. La deuxième étape concerne l'assignation des patients aux physiothérapeutes, en prenant en compte l'équité, dans le but de créer des plannings équilibrés pour chaque physiothérapeute. Pérez *et al.* (2011) ont proposé deux algorithmes pour une planification efficace des ressources et des patients en médecine nucléaire. Leur approche a été évaluée via une étude de cas qui permet de la comparer avec les méthodes utilisées dans une clinique de médecine nucléaire. Dans une étude antérieure (Pérez *et al.*, 2013), ils sont passé à la version en ligne stochastique des algorithmes de planification des patients et des ressources. Huang *et al.* (2022) ont présenté des algorithmes basés sur l'optimisation pour maximiser l'efficacité de l'utilisation des ressources de surveillance neurophysiologique. Ils ont planifié les cas chirurgicaux et les technologues en réduisant les heures supplémentaires et le temps de sous-utilisation.

2.5 La replanification en santé

L'incertitude est une préoccupation très importante dans la planification de la production, des perturbations peuvent être engendrées par ce phénomène, de ce fait, il est primordial de la prendre en compte pour assurer la faisabilité des plannings réalisés (Li et Ierapetritou, 2008). Les événements inattendus dans les processus de production peuvent causer des temps d'arrêt de fabrication considérables, d'où la nécessité de la replanification des plannings affectés par les changements afin de maintenir la bonne performance du système de production (Uhlmann et Frazzon, 2018).

La replanification est le processus de mettre à jour les calendriers de production existant en raison de perturbations ou changements, y compris l'occurrence des nouvelles tâches, la panne des machines, etc. (Vieira *et al.*, 2003). Uhlmann et Frazzon (2018) ont confirmé que la replanification des plannings consiste à déterminer les stratégies à utiliser et le moment pour réagir. Les stratégies utilisées sont réparties en deux : la réparation des plannings en appliquant des ajustements sur le planning initial, et la replanification complète en générant de nouveaux plannings. La replanification peut être réalisée périodiquement, à l'occurrence des événements, ou bien d'une manière hybride qui combine les deux premières méthodes.

De même, le processus de planification en santé confronte l'apparition des événements imprévus qui peuvent impacter la performance, comme le retard des patients, la non-disponibilité des ressources, et la panne des équipements (Mageshwari et Kanaga, 2012). Ces changements liés à un tel environnement dynamique imposent le recours à la replanification des plannings générés initialement. Dans la littérature, il existe des études qui concernent la replanification du personnel médical ainsi que la replanification des patients.

Gross *et al.* (2018) ont présenté la première étude qui traite le problème de la replanification des médecins en cas d'absence, en développant un modèle de programmation mixte en nombres entiers. De plus, Clark *et al.* (2015) ont réalisé une revue de littérature comportant les travaux liés à la replanification des horaires des infirmiers. Une étude récente de Long *et al.* (2022) concerne le problème de la replanification des infirmières avec des disponibilités variables et incertaines. Les auteurs ont développé un modèle de programmation stochastique basé sur des données historiques, et un modèle robuste pour affronter le problème d'insuffisance de données. Le modèle robuste fournit des solutions plus fiables et flexibles. En outre, en se basant sur une étude de cas réel, les auteurs ont confirmé que les modèles développés sont efficaces et permettent de réduire les coûts hospitaliers.

La replanification des patients en chimiothérapie est très répandue dans les centres d'oncologie, causée par des complications de traitement, des conflits d'horaire des patients, et des raisons administratives de l'établissement (Gunasekaran *et al.*, 2020). Cette étude concerne les patients atteints par le cancer de sein, et confirme que 69 % des patients considérés dans la recherche subissent des changements dans leur planning. Condotta et Shakhlevich (2014) et Hooshangi-Tabrizi *et al.* (2020) ont traité le problème de planification des rendez-vous des patients en chimiothérapie, ils ont suivi en une approche de replanification quotidienne qui permet une meilleure allocation des ressources en cas de la survenance des événements imprévus.

Stuart et Kozan (2012) ont développé un modèle de replanification pour une salle d'opération, et ils l'ont résolu à l'aide d'une heuristique. L'objectif est de maximiser le nombre pondéré d'interventions chirurgicales prévues à effectuer à temps. Une approche de planification prédictive-réactive est présentée par Wang *et al.* (2015) pour traiter les chirurgies en urgence aléatoire tout en optimisant l'efficacité et la stabilité. Vali-Siar *et al.* (2018) ont présenté un problème de planification et de replanification des salles d'opération multi-périodes et multi-ressources avec des patients urgents (électifs) et semi-urgents (semi-électifs). Ce problème a été modélisé avec une programmation linéaire à nombres entiers mixtes selon une approche d'horizon glissant. Les auteurs visent à minimiser les retards, les temps d'inactivité et les heures supplémentaires. Dans un premier temps, un planning initial est établi. Ensuite, une replanification est effectuée après l'arrivée d'un patient semi-urgent. La replanification a été aussi utilisée dans les services de bloc opératoire en raison d'une libération du bloc pour réaffecter les créneaux horaires disponibles à d'autres chirurgiens afin de répondre à la demande réelle et de maximiser l'efficacité du bloc opératoire (Akbarzadeh *et al.*, 2019). Davarian et Behnamian (2022) ont traité le problème de la planification et la replanification des salles d'opération due à l'annulation des patients. Dans cette étude, chaque patient passe par deux étapes à l'hôpital. Il commence par la salle d'opération, par la suite, il est transféré au service de réveil après la chirurgie. Une approche d'optimisation robuste a été proposée pour traiter les incertitudes sur la durée de la chirurgie et le nombre de lits en salle de réveil. Pour résoudre le problème, l'algorithme génétique a été utilisé.

En nous basant sur la présente revue de littérature, nous concluons qu'il n'existe pas des sujets de recherche qui étudient la replanification des patients en radiothérapie, notamment suite à un manque de ressource tel que la panne des machines.

CHAPITRE 3 MÉTHODOLOGIE

Cette thèse contient des applications réelles des sujets de recherche en santé grâce aux collaborations avec le CICL et le CHUM. Notre objectif est d'améliorer le système de planification des rendez-vous des patients et des horaires des technologues à travers l'extraction de l'information utile des données massives et l'optimisation des processus. Nous avons réalisé des analyses statistiques des données, en outre, nous avons utilisé des techniques d'apprentissage machine, ainsi que de la recherche opérationnelle.

3.1 Prédiction du temps de service des patients

La stratégie de planification de rendez-vous sans bloc permet d'attribuer des créneaux de différentes durées aux patients selon le traitement requis, contrairement à la stratégie de planification avec bloc qui assigne des rendez-vous de longueur uniforme à tous les patients (Conforti *et al.*, 2010). Bien que plusieurs centres de radiothérapie appliquent la stratégie avec bloc, celle sans bloc représente une grande opportunité d'amélioration en tirant profit de temps total excédent des rendez-vous dont la durée de traitement est moins que prévue.

Dans le premier article présenté dans cette thèse, nous reconstruisons de nouvelles grilles des rendez-vous des patients en radiothérapie en redéfinissant la longueur des créneaux horaires. Nous avons accès aux données réelles du CICL. Nous utilisons un modèle de prédiction du temps de traitement à l'aide des techniques d'apprentissage machine en suivant le processus de CRISP-DM composé de six phases :

1. Objectif de la recherche : Cette phase du processus consiste à déterminer les objectifs. Nous visons à prédire le temps de traitement en radiothérapie des patients du CICL.

2. Compréhension des données : Nous collectons et nous évaluons les données. Elles ne contiennent que quatre attributs (catégorie du cancer, plan de traitement, statut, salle de traitement) qui sont en lien avec le temps de service. Nous réalisons une analyse descriptive pour étudier la variation du temps de service par rapport à chaque attribut de prédiction.

3. Préparation des données : Nous nettoyons les données à partir des données brutes initiales. En outre, nous faisons un test de sélection des attributs les plus importants. La variable de sortie devrait être de nature catégorique afin de l'utiliser dans la conception des grilles des rendez-vous. Nous faisons transformons le temps de service en variable catégorique dans le cas des techniques de classification à l'aide des classe de temps de différentes durées.

4. Modélisation : Nous appliquons quatre techniques d'apprentissage machine (la régression linéaire, MARS, les réseaux de neurones, l'arbre de classification et de régression).

5. Évaluation : La qualité des modèles élaborés est évaluée en se basant sur leur précision.

6. Déploiement : Nous exploitons le modèle de prédiction dans la réalisation des nouvelles grilles des rendez-vous.

3.2 Optimisation intégrée de la planification des grilles des rendez-vous et des horaires des technologues

Dans la plupart des centres d'imagerie médicale, la planification des grilles des rendez-vous et des horaires des technologues est faite d'une manière séparée. Or, plusieurs examens médicaux peuvent être effectués par un ou deux technologues. Une amélioration de la performance peut être réalisée par la combinaison de ces deux éléments du système de prise des rendez-vous en considérant le changement apporté par la modification du nombre des technologues lors de l'exécution des examens médicaux.

Le deuxième article de la thèse concerne la planification des grilles des rendez-vous et des horaires des technologues au centre d'imagerie médicale. Pour ce faire, nous présentons deux versions du modèle mathématique proposé : séquentiel et intégré. La première version traite les grilles des rendez-vous et les horaires des technologues d'une manière séquentielle ; or, la deuxième combine les deux. L'objectif est de décider le nombre, le type et la séquence des examens affectés à chaque machine, ainsi que les heures de début et de fin de quart, les pauses, les jours de congé et la machine assigné pour chaque technologue. Le modèle d'optimisation considère l'allocation optimale des technologues pour maximiser le nombre de patients servis. Donc, nous réalisons une analyse statistique basée sur les données réelles pour étudier l'impact du nombre des technologues alloués sur le nombre total des patients.

3.3 Optimisation hiérarchique pour la replanification des patients

L'incertitude est présente d'une manière considérable dans la planification des rendez-vous des patients, elle peut être liée à la variabilité des temps de service, l'arrivée des patients, la disponibilité des ressources, etc. Le processus de traitement en radiothérapie est basé principalement sur une ressource primordiale qui est l'accélérateur linéaire. La panne de cette machine peut engendrer des coûts importants. Afin de minimiser ces coûts, les planificateurs font des décisions rapides, leur impact n'est pas évalué en prenant compte de la performance globale, ils essayent d'éviter principalement la situation la plus pire.

Dans le troisième article de la thèse, nous présentons la replanification des rendez-vous des patients en radiothérapie. Nous déterminons toutes les séquences des priorités de trois décisions de replanification qui sont : retarder un patient, surréserver un patient, utiliser le temps supplémentaire pour céduler un patient. Nous proposons une approche multiobjectif qui minimise le nombre de patients retardés, l'attente, et le temps supplémentaire. Les objectifs sont traités un par un d'une manière consécutive. Nous proposons une heuristique qui suit le cadre de replanification que nous avons défini. En outre, nous développons un modèle mathématique en appliquant l'optimisation hiérarchique. Il s'agit de commencer à minimiser un objectif donné en considérant toutes les contraintes du problème, par la suite, nous passons à optimiser un autre objectif en ajoutant la valeur maximale de l'objectif précédent dans les contraintes.

CHAPITRE 4 ARTICLE 1: PATIENT SCHEDULING BASED ON A SERVICE-TIME PREDICTION MODEL: A DATA-DRIVEN STUDY FOR A RADIOTHERAPY CENTER

D. Ben Tayeb, N. Lahrichi et L-M. Rousseau

Cet article a été publié dans Health Care Management Science, volume 22, pages 768–782 (11 Octobre 2018).

4.1 Abstract

With the growth of the population, access to medical care is in high demand, and queues are becoming longer. The situation is more critical when it concerns serious diseases such as cancer. The primary problem is inefficient management of patients rather than a lack of resources. In this work, we collaborate with the Centre Intégré de Cancérologie de Laval (CICL). We present a data-driven study based on a nonblock approach to patient appointment scheduling. We use data mining and regression methods to develop a prediction model for radiotherapy treatment duration. The best model is constructed by a classification and regression tree; its accuracy is 84%. Based on the predicted duration, we design new workday divisions, which are evaluated with various patient sequencing rules. The results show that with our approach, 40 additional patients are treated daily in the cancer center, and a considerable improvement is noticed in patient waiting times and technologist overtime.

keywords: Patient scheduling, Data-driven approach, Prediction models, Nonblock scheduling, Grid design, Sequencing rules

4.2 Introduction

Nearly half of all Canadians will be diagnosed with cancer during their lifetime. Cancer is the leading cause of death in Canada (Société canadienne du cancer, 2017). These statistics indicate that it is vital to ensure timely access to medical care. However, given the continued growth in the number of cancer patients, an imbalance between appointment demand and treatment capacity has arisen. Therefore, waiting times are becoming longer. This leads to patient dissatisfaction and higher costs for clinics. The critical factor is usually suboptimal patient scheduling rather than limited resources. To face these challenges, cancer centers must better manage patient appointments. In this paper, we a present data-driven study that develops decision support tools based on data mining to improve patient scheduling. We collaborate with the department of radiotherapy at the Centre Intégré de Cancérologie de Laval (CICL).

The decisions involved in planning an outpatient appointment system can be classified into three categories: strategic, tactical, and operational (Ahmadi-Javid *et al.*, 2017). We consider the tactical and operational levels separately and sequentially. The tactical level includes the development of a more reliable appointment interval, which directly impacts the number of appointments in a session. The operational level involves determining the patient appointment times according to a given sequencing rule.

There are two appointment scheduling strategies: block and nonblock systems (Conforti *et al.*, 2010). The block system divides the day into a fixed number of slots with the same duration, whereas the nonblock system allows appointment intervals of different durations. The CICL uses a block policy. Radiotherapy appointment scheduling is complex, since the treatment is divided into several sessions that occur on successive days, and the CICL requires all the treatments for a patient to take place at the same time of day and in the same room. Therefore, it is simpler to apply the block strategy. However, most of the allocated slots are not respected; the service time differs from one patient to another.

Appointment management systems that allocate uniform slots assume that patients are similar, and all the treatments have a given average duration that determines the length of the appointment interval. In reality, the patients differ in terms of disease category and management, and even the features of the visits vary (the first treatment or the last, the patient arrives from the emergency room or the hospital, etc.). Several factors affect the patient service time, so it is sensible to allow considerable variation in these durations. In the block scheduling system, treatments that end early compensate for others that exceed the expected time. However, the lack of control of the service-time variability generates indirect costs related to capacity underutilization, overtime, and waiting time.

This paper develops a data-driven approach to efficient appointment scheduling based on the nonblock strategy. The method has two phases. The first phase (Section 4.4) predicts the patient service time based on the treatment characteristics. We apply data mining and regression tools to extract information from medical data. The main goal is to classify patients according to their treatment durations. The prediction algorithm must be very accurate in order to lead to effective patient scheduling. The second phase (Section 4.5) is a patient appointment system based on the prediction model. We design the appointment grid and determine patient management and sequencing rules. Reliable patient service times lead to better scheduling; we aim to maximize the outpatient clinic utilization and minimize the patient waiting time and the technologist overtime.

To the best of our knowledge, this work is the first radiotherapy application that uses data mining and regression methods to define patient appointment durations and proceeds to develop a more efficient patient schedule based on the nonblock strategy. It is a data-driven study that uses real CICL Radiotherapy data. The steps are as follows: current scheduling strategy analysis; extraction of actionable models; development of patient schedule; determination of patient sequencing and operational management rules; evaluation; and validation. Compared to the current scheduling system, our approach gives good results in terms of waiting time, overtime, and the number of patients seen per day; the improvement reaches 30%.

The remainder of the paper is organized as follows. Section 2 summarizes related work and gives the problem statement. Section 3 discusses the development of our prediction model for the service time. Section 4 discusses the development of the appointment scheduling system and the results. Finally, Section 5 provides concluding remarks.

4.3 Problem statement and related work

The CICL workday is split into a fixed number of 20 min appointments. However, the actual treatment time may vary, so for many patients the appointment times are not respected. They can either arrive early or risk waiting a long time. Figure 4.1 shows that the median waiting times in September 2017 ranged from 0 to 8 min, with the largest variability on day 28.

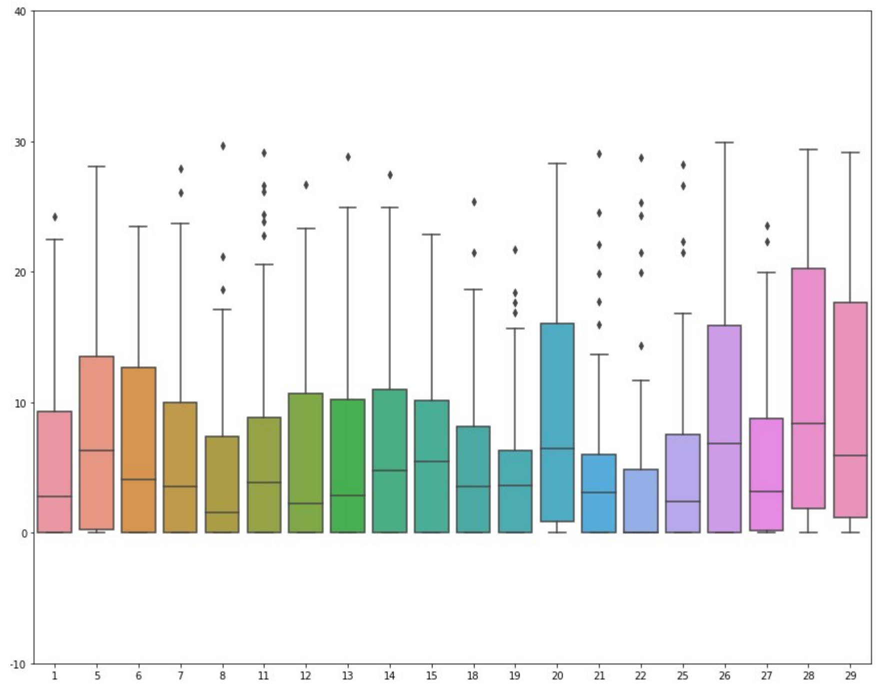

Figure 4.1 Patient waiting times in September 2017.

The CICL has four active rooms with linear accelerators, and 32 patients can be treated per day in each room. However, the demand may exceed this limit. Also, technologists may finish early or work overtime.

Our goal is to develop an efficient patient scheduling method. Outpatient appointment scheduling has been widely studied, starting with the pioneering work of (Bailey, 1952). Three reviews (Cayirli et Veral, 2003; Gupta et Denton, 2008; Ahmadi-Javid *et al.*, 2017) provide a global overview of developments. Cayirli et Veral (2003) present various formulations based on mathematical programming, simulation, and queuing theory. Gupta et Denton (2008) define different types of health care systems, focusing on the factors that complicate appointment scheduling. Ahmadi-Javid *et al.* (2017) classify the scheduling decisions into three classes: strategic, tactical, and operational.

Cayirli *et al.* (2006) show that appointment scheduling systems can be improved by classifying the patients according to, e.g., the type of procedure required or the variability of the service time. They differentiate between new and returning patients. Based on this classification, they define six patient-sequencing rules. They simulate the effect of these rules in combination with seven scheduling rules, based on factors such as the number of patients per session and the probability of no-shows. The results show that the sequencing rules have a considerable impact on the performance of appointment management systems in outpatient clinics. Approaches that combine sequencing and appointment timing are considered complex. Most papers assume that the sequence is known, or they apply heuristic sequencing rules such as first-come-first-served. The most popular rule, which has performed well in several studies (e.g., (Denton *et al.*, 2007; Gupta, 2007)), is smallest-variance-first (SVF), which orders patients by increasing order of service-time variance.

The existence of huge medical databases has encouraged researchers to carry out data-driven studies. Bakker et Tsui (2017) develop a data-driven approach to dynamic resource allocation for patient scheduling. They perform a discrete-event simulation with the empirical data to compare their method to the traditional cyclic schedule and to the resource calendar at the hospital. Huang et Bach (2016) perform a data-driven study to determine an appointment target lead time policy.

Mandelbaum *et al.* (2020) propose an infinite-server approach for appointment scheduling and sequencing problems. It outperforms a data-based robust-optimization algorithm that is near-optimal for the single-server problem. Their tests use real data from the cancer center's infusion units, and they decrease waiting time and overtime by 30%. Kim *et al.* (2018) aim to understand the existing appointment schedules. Analyzing data from an endocrinology clinic, they construct a high-fidelity simulation model of the stochastic arrival process.

The data stored by health facilities is too large and heterogeneous to be processed by traditional statistical methods. Therefore, it is necessary to use powerful tools such as data mining techniques to extract significant information (Koh *et al.*, 2011). We are interested in data-based studies that use prediction models constructed with data mining techniques.

Most data mining techniques provide information that classifies patients in terms of no-shows or hospital readmission or appointment length. For the prediction of hospital readmissions, Golmohammadi et Radnia (2016) state that their work differs from previous studies (Lagoe *et al.*, 2001; Billings *et al.*, 2006; Donnan *et al.*, 2008; van Walraven *et al.*, 2012) in the number of data mining techniques applied. Instead of a single model based on logistic regression, they use neural networks, classification and regression (CR) tree, and chi-squared automatic interaction detection. The models have an overall accuracy above 80%. Simiarly, Braga *et al.* (2014) construct several models for readmission prediction by exploiting the support vector machine, the naive Bayesian classifier, and decision trees. The best results are obtained by the naive Bayesian classifier, with an accuracy of 98.91%.

Articles that study no-shows indicate that estimating the no-show probability has a positive impact on the overbooking. Lotfi et Torres (2014) compare four decision tree techniques and conclude that the CR tree is the most powerful; it also works better than Bayesian networks and neural networks. Based on the no-show probabilities extracted from the tree, five scheduling policies are simulated to evaluate the impact of variation in overbooking levels. Huang et Hanauer (2014) propose a logistic regression model to predict the probability of no-shows. It is also used to calculate the no-show threshold, which is used to determine the status of the patient (missing the appointment or not) in order to overbook a patient in the case of no-show. This approach is compared by simulation with two standard overbooking policies. It gives the best results in terms of reducing patient waiting times, physician idle time, and overtime.

Other authors use hybrid methods. Glowacka *et al.* (2009) apply a hybrid data mining/simulation approach. They use association rules to predict the probability of no-shows. The simulation is used to evaluate the chosen rules with four scheduling methods, integrating the no-show probability of each patient, to optimize the number of patients scheduled. Alaeddini *et al.* (2011) present a probabilistic hybrid model based on logistic regression and Bayesian inference, which is used to update the no-show probability. This technique is compared to other methods, including time series, decision trees, and logistic regression. More recently, Harris *et al.* (2016) have developed a new model of no-show prediction, combining a regression model and a functional approximation based on the sum of exponential functions. It is compared to other methods of predicting binary data, such as logistic regression and the CR tree.

Huang et Marcak (2013) exploit a decision tree to propose a schedule based on time slots that are a multiplication of 15 min instead of 30 min. They apply a decision tree to classify patients based on their characteristics, and the results are used to assign an appropriate time interval. The approach increases radiographer utilization and patient access.

We use a data-driven approach to develop an appointment schedule based on a nonblock policy. We apply various data mining and regression tools to the patient treatment features to choose the best service-time prediction model. This model determines an adequate service time for each patient, and this is used to construct new time grids. Several grids are evaluated with different patient sequencing and management rules to select the best appointment system.

4.4 Development of the prediction model

In this section, we develop a tool for predicting the patient treatment times. The service or treatment time represents the total time during which the technologists interact with and treat a patient during a given session. The prediction model must be very accurate in order to lead to efficient patient scheduling. Thus, several prediction methods are compared. We work with real data from CICL Radiotherapy. This study follows the CRISP-DM process (CRoss Industry Standard Process for Data Mining).

4.4.1 Data mining process

Larose et Larose (2014) present data mining as a well-structured standard process. They claim that every data mining project follows the CRISP-DM, which has six phases (see Fig. 4.2):

1. Business understanding: The first phase expresses objectives and prepares a strategy to achieve them.
2. Data understanding: This phase collects the data and evaluates their quality. It also identifies interesting subsets that may contain actionable models.
3. Data preparation: This phase cleans the raw data, selects attributes to analyze, and performs any necessary variable transformations.

4. Modeling: Several modeling techniques can be used. During this phase, the data may be adapted for a specific data mining method.

5. Evaluation: This phase evaluates the quality and effectiveness of the models from the previous phase.

6. Deployment: The project does not end with the creation of the model. It must be further developed and adapted according to the customer's needs.

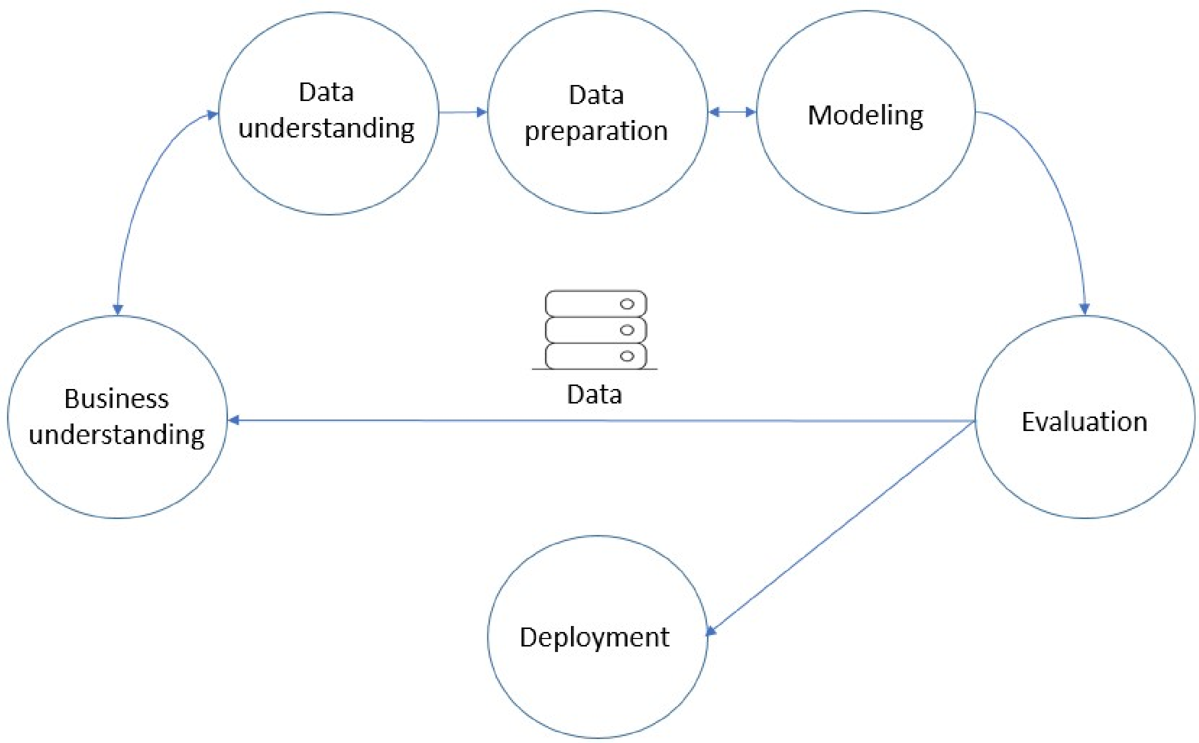

Figure 4.2 CRoss Industry Standard Process for Data Mining (CRISP-DM).

4.4.2 Application of the CRISP-DM to CICL service time prediction

a) Business understanding

The goal of this study is to predict the service time for radiotherapy patients based on their treatment characteristics, and to identify the important variables that provide this information.

b) Data understanding

The data correspond to treatments performed at the CICL from 2012 to 2016. They include information related to the patient's visit (date of appointment, start time and end time, etc.), and treatment features (category, room, etc.).

At this step, it is essential to perform a descriptive analysis to gain a clear idea of the current state. We start by verifying the accuracy of the allocated service time. Figure 4.3 shows that the average treatment time ranges from 2 to 22 min, but all the appointment intervals are the same (20 min). Adjusting the durations could lead to better patient scheduling.

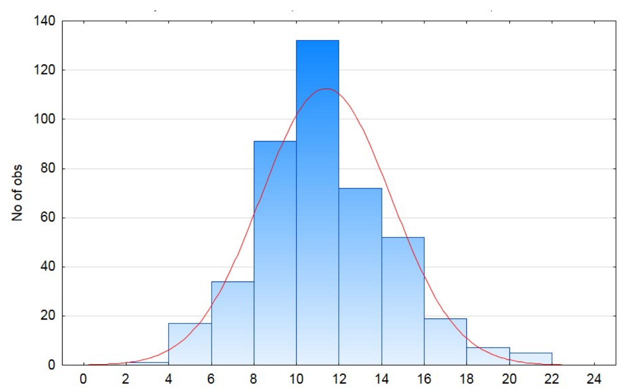

Figure 4.3 Average treatment times in CICL.

We therefore analyze the treatment attributes (cancer category, care plan, status, and treatment room) and their impact on the treatment duration. Figure 4.4 indicates that the main cancer types are breast and prostate cancer. Figure 4.5 shows that most cancers have a treatment time between 10 and 12 min. Cancers of the vulva and anal canal require more time, whereas seminoma, brain, and sinus cancer are treated more quickly. Figure 4.6 shows that these cancers also have the shortest average treatment time by care plan. Figure 4.7 shows the average treatment time by status; this combines information about the patient and the treatment session. The status indicates, e.g., if the patient is hospitalized (H), if there is a change of treatment plan (PS), and if the session is the first (DC) or last (FC). The first treatment usually takes more time. Figure 4.8 indicates that the average treatment time is almost the same for all the rooms.

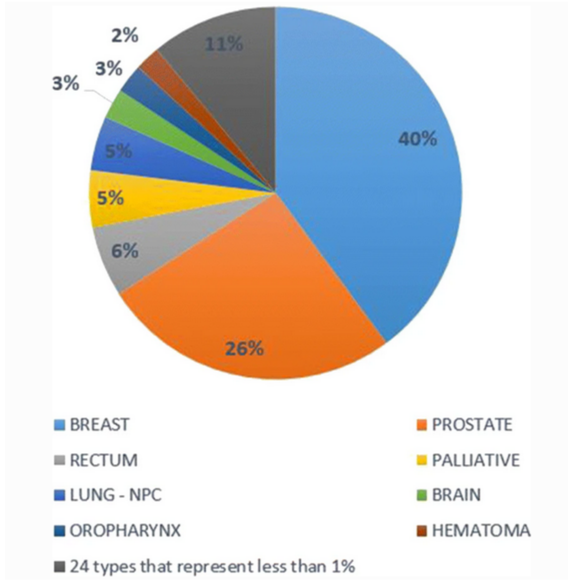

Figure 4.4 Cancer categories in CICL.

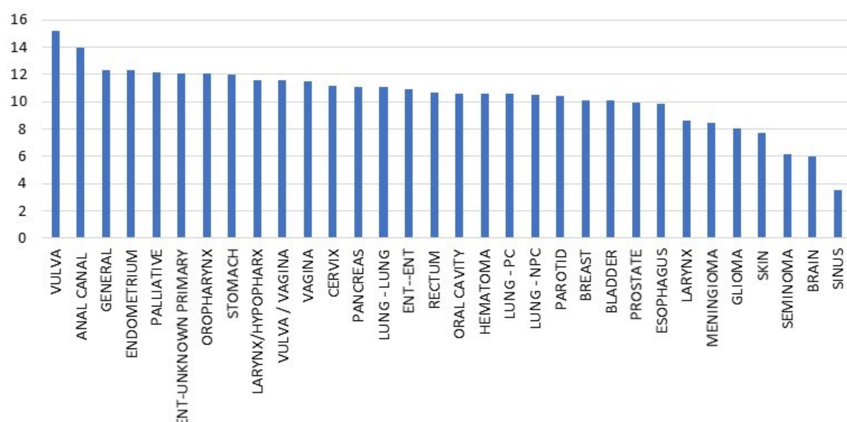

Figure 4.5 Average treatment times by cancer category.

c) Data preparation

This phase consists of cleaning the data, adding new attributes, and preparing and selecting the attributes.

Data cleaning We start by removing unnecessary information and outliers or incorrect data. Negative or excessively long periods are also removed. This eliminates less than 3% of the data.

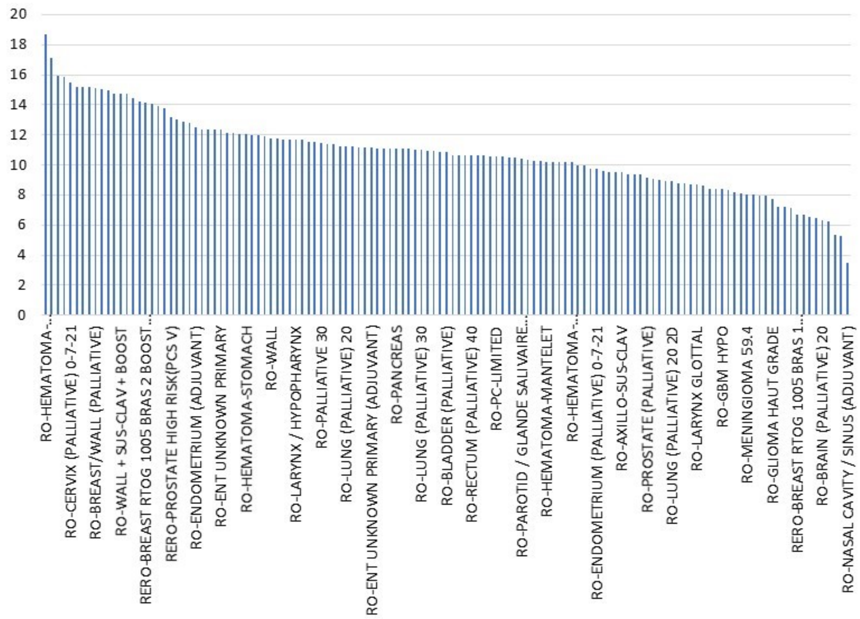

Figure 4.6 Average treatment times by care plan.

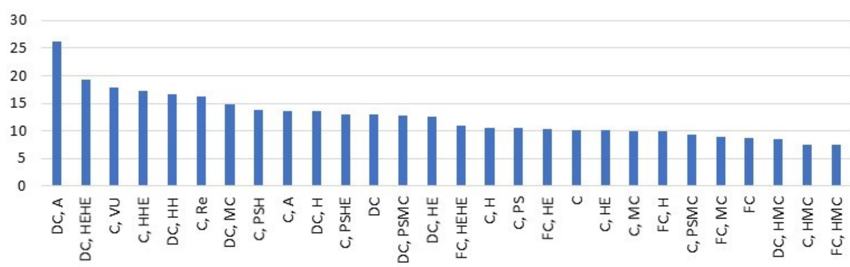

Figure 4.7 Average treatment times by status.

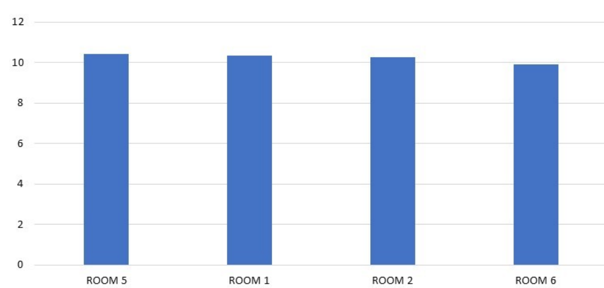

Figure 4.8 Average treatment times by room.

Attribute addition We calculate the treatment duration, which starts at the moment that the patient enters the treatment room, and then find the average duration.

Attribute preparation Because we plan to design a standard appointment grid, we convert the continuous output into a categorical variable with the following classes: 5, 10, 15, 20, and 25 min.

Attribute selection This step identifies the important attributes. We used the chi-square test χ^2, which is a statistical test of independence between variables; Table 4.1 summarizes the results. The p-values indicate that the most significant variables are cancer category, care plan, and status. We therefore consider only these three variables.

Table 4.1: Results of the chi-square test

Variable	P-value
Cancer category	0.0000
Care plan	0.0000
Status	0.0000
Appointment room	0.1004

d) Modeling

In this phase, we develop prediction models for the treatment duration, exploiting data mining and regression tools and using the Statistica software.

General linear model This statistical technique models the relationship between the explanatory variables, which can be continuous or categorical, and the output. The model is built using a linear predictor function:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \epsilon \tag{4.1}$$

$$\begin{aligned} \text{where} \quad & y && : \textit{dependent variable;} \\ & x_i && : \textit{independent variable;} \\ & \beta_i && : \textit{parameters to estimate;} \\ & \epsilon && : \textit{normally distributed error.} \end{aligned}$$

This model predicts a continuous output. The results are good; see Table 4.2. The coefficient of determination R^2 shows that 80% of the data variability is explained by the model. We next categorize the variable predicted by this model. Figure 4.9 compares the predicted and observed classes. It shows that the two largest classes (10 and 15 min) are generally well predicted.

Table 4.2: Results of the general linear model

Dependent variable	R^2	Sum of squared errors	Mean square error	P-value
Treatment time	0.809	758.044	2.669	0.00

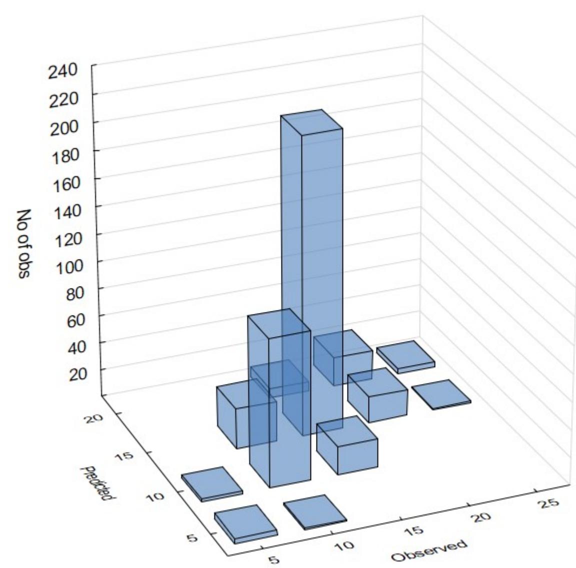

Figure 4.9 General linear model classification graph.

Multivariate adaptive regression splines (MARS) This algorithm performs piecewise linear regression by partitioning the data set based on optimal nodes and assigning each subset an equation or a classification. The classification (Fig. 4.10) shows that only the 15 min class is well predicted.

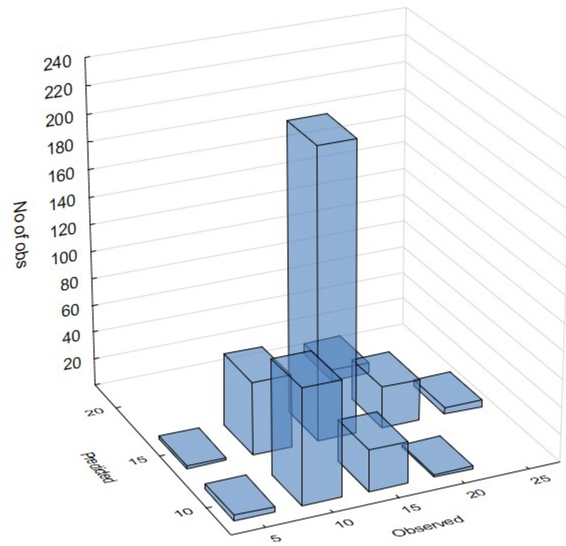

Figure 4.10 Mars classification graph.

Artificial neural networks This technique simulates the operation of biological neural networks in the human brain. Artificial neural networks contain a set of interconnected elements called neurons. The network consists of three layers. The first is the input layer, containing the nodes that represent the independent variables. The last is the output layer, containing the dependent variable. Between these layers there is a hidden layer, in which the nodes are not observed but calculated based on the input variables. The nodes of the network are connected by arcs with different weights. The neural network construction algorithm is adaptive: it changes its structure and adjusts the weights to minimize the error.

The resulting network has six neurons in the hidden layer, and the activation function for the hidden and output layers is the logistic function (Table 4.3). The classification graph (Fig. 4.11) shows that only two classes (10 and 15 min) are well predicted.

Table 4.3: Results for Neural Networks

Network name	Training performance	Test performance	Validation performance	Hidden activation	Output activation
MLP 179-6-5	80.974	75	76.562	Logistic	Logistic

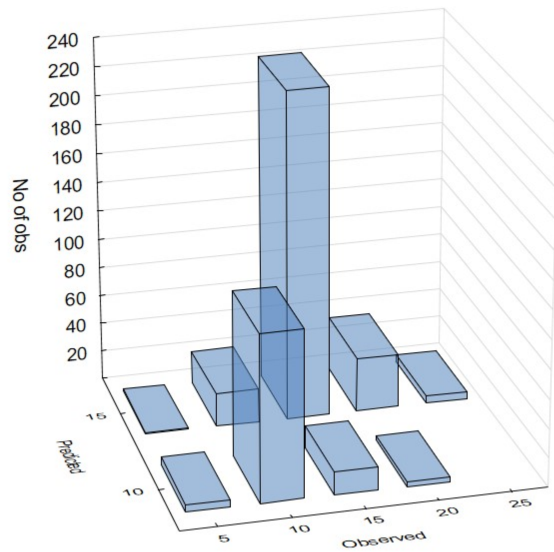

Figure 4.11 Neural network classification graph.

CR tree This is constructed iteratively, by separating the population at each stage into two groups, in order to maximize the purity of the nodes. Each nonterminal node in the tree represents a test on an attribute, and each branch signifies the result of a test. A class label or an average is assigned to each leaf of the tree. This tree gives satisfactory results. In addition, it is easy to interpret because it is not very deep (Fig. 4.12). From the classification graph (Fig. 4.13), we deduce that most of the predicted values are correct. The important variables are status and care plan; however, the category is used in only one test in the tree.

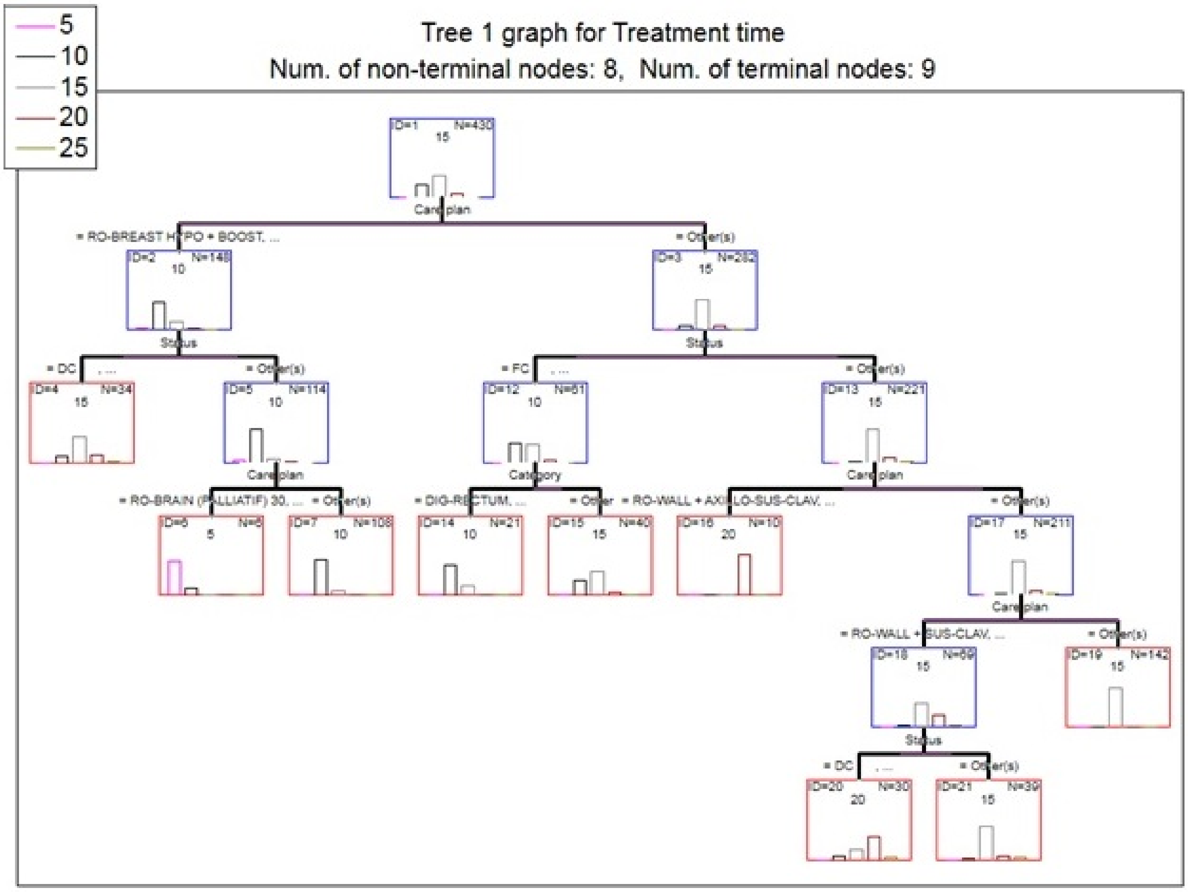

Figure 4.12 Results of classification and regression tree.

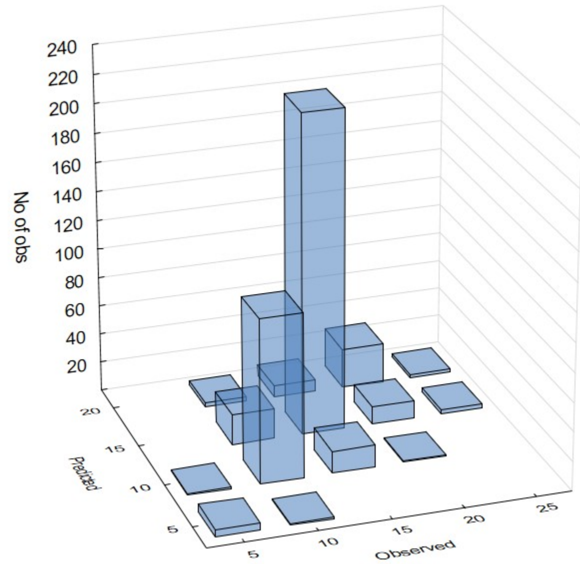

Figure 4.13 Classification and regression tree classification graph.

e) Evaluation

In this section we evaluate the performance of the models. The evaluation criterion depends on the type of the output variable. Our dependent variable is categorical, so we compare the models based on their accuracy:

$$Accuracy = \frac{Number\ of\ correct\ predicted\ values}{Total\ number\ of\ observations}. \tag{4.2}$$

We deduce from Table 4.4 that the best model is based on the CR tree.

We perform an additional analysis of the prediction model from the CR tree: we examine the prediction error for each class. Table 4.5 shows that the dispersion of the absolute prediction

Table 4.4: Performance of prediction models

Prediction method	Accuracy
CR tree	84%
General linear model	81%
Artificial neural networks	76%
MARS	71%

error is almost the same for all the classes and not very large; moreover, the mean square error for the 5 min and 10 min classes is minimal compared to that for the 15 min and 20 min classes. The boxplots of the residuals between the observed and predicted values confirm this. Figure 4.14 indicates that the median residuals for the 5 min and 10 min classes are about -1 and -2 min respectively; however, for the larger classes the residuals are about -5 and -6 min respectively.

Table 4.5: Prediction error

Class	Mean square error	Standard deviation
5 min	9.33	2.32
10 min	13.52	2.17
15 min	36.3	2.45
20 min	48.57	2.97

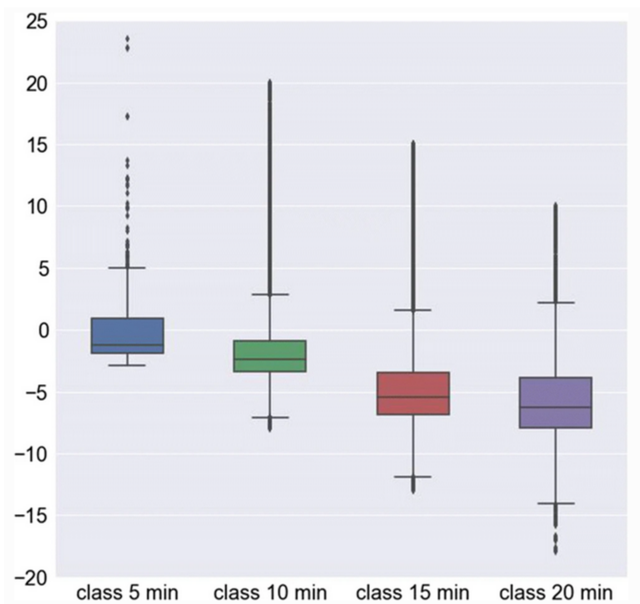

Figure 4.14 Residuals between observed and predicted treatment times.

f) Deployment

The model based on the CR tree is characterized by its simplicity and good performance. In Section 4.5, we study the impact of this decision tool on the appointment scheduling. We first design the appointment grid and then determine the management and sequencing of the patients.

For the grid, we use the prediction model to determine the number of each duration class. However, we must also consider the interpatient duration. This is the time allowed between two successive patients to prepare the room and the material. Historical data analysis indicates that this time is about 5 min. We set the interpatient duration to 0 or 5 min. We allow it to be 0 for the 15 min and 20 min classes. In these cases, the treatment time has a median residual of about -5 min; we interpet this as an interpatient duration included in the expected service time.

4.5 Appointment patient scheduling in the CICL

Using our prediction model, we construct schedule grids according to the nonblock strategy. We compare various sequencing and management rules. The evaluation is based on the number of patients seen per day, the waiting time, and the technologist overtime.

4.5.1 Grid design

The grid design has three steps: 1) Patient class distribution; 2) Interpatient duration estimation; and 3) Workday division.

Patient class distribution We carry out an analysis of data from 2017, after having applied our prediction algorithm. We calculate the average number of each treatment duration to determine the daily patient class distribution (see Table 4.6). The 15 min class is the most frequent, followed by 10 min and 20 min. The 5 min class is not used, so we will not consider this duration in the grid.

Table 4.6: Average number of each class per day during 2017

Class	Average number
5 min	0
10 min	48
15 min	77
20 min	4

Interpatient duration estimation This duration varies considerably, with a median around 5 min. We set this duration to 5 min or to 0 if it is covered by the prediction residual.

Workday division We consider various divisions of the workday. Each division defines the average number of each class, the interpatient duration, and a slot sequence. Figure 4.15 illustrates four divisions. The first grid respects the number of time intervals in Table 6 and sets the interpatient duration to 5 min; there are 12 slots of 15 min, 22 slots of 20 min, and 1 slot of 25 min. In the second grid, the interpatient duration is set to 5 min for the 10 min class, and to differentiate this class from the 15 min class, the former time slots are colored green. In the third grid, the green and red 15 min intervals are alternated. In the fourth grid, the sequence of the intervals is chosen randomly.

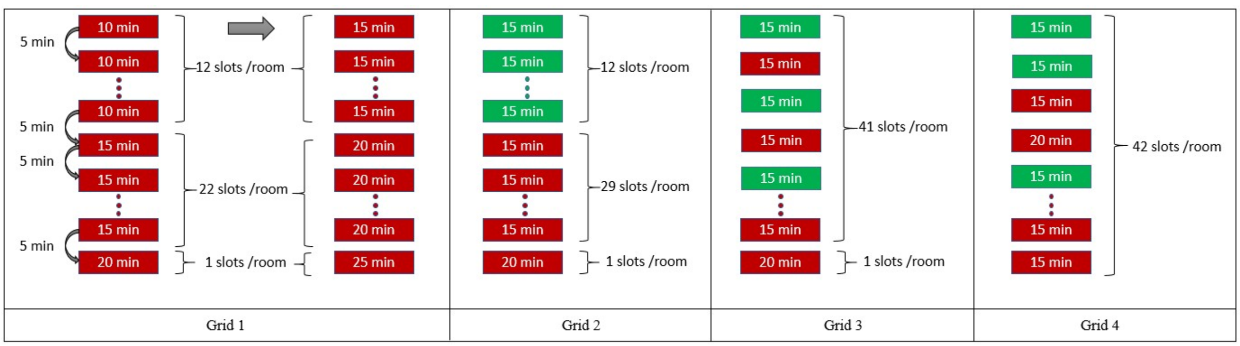

Figure 4.15 Divisions of the CICL working day.

4.5.2 Sequencing rules

We apply the following sequencing rules: SVF, smallest-mean-first (SMF), and an assignment without rules. For the first two rules, the patients are sorted in ascending order of the variance (or mean) of the treatment times, and the allocation of the slots respects this order. For the third grid, we test two more scenarios. These alternate between the small and the large variance (or mean) of the treatment duration.

4.5.3 Operational management rules

There are two ways to manage the treatment start: either the patients must wait until the scheduled start time, or they can be treated early if the machine and the technologists are free. The CICL adopts the second strategy, and most of the patients arrive before their appointment times. The choice of strategy affects the waiting time and technologist overtime. We test both methods and compare key performance indicators.

4.5.4 Performance indicators

The performance indicators are: 1) Patient waiting time, 2) Technologist overtime, and 3) Number of patients seen per day.

Patient waiting time This is divided into indirect waiting time and direct waiting time. The first is the difference between the date of the appointment request and the date of the consultation. The second is the positive difference between the treatment start time and the maximum of the arrival time and the appointment time (Cayirli et Veral, 2003). We consider only the direct waiting time.

Technologist overtime This is the positive gap between the treatment completion time for the last patient and the expected end of the workday (Cayirli et Veral, 2003).

Number of patients seen per day The maximum number of patients per day depends on the number of slots that can be scheduled.

In this study, an efficient schedule increases the number of patients per day and decreases their waiting time and the technologist overtime.

4.5.5 Grid evaluation

To illustrate our approach, we take the data for September 2017. The treatment times are generated by Monte Carlo simulation using the residual between the observed and predicted treatment times.

Figures 4.16, 4.17, and 4.18 illustrate the residuals grouped by cancer category and by duration class. The median of this variable is between -0.8 and -11 min. The greatest variations in the values occur in the skin category of the 10 min class, in the palliative category of the 20 min class, and in the vulva/vagina, sinus, and skin categories of the 15 min class.

We apply the Monte Carlo technique to a subset of the residuals, to arbitrarily produce new treatment times by adding the duration class to this random variable.

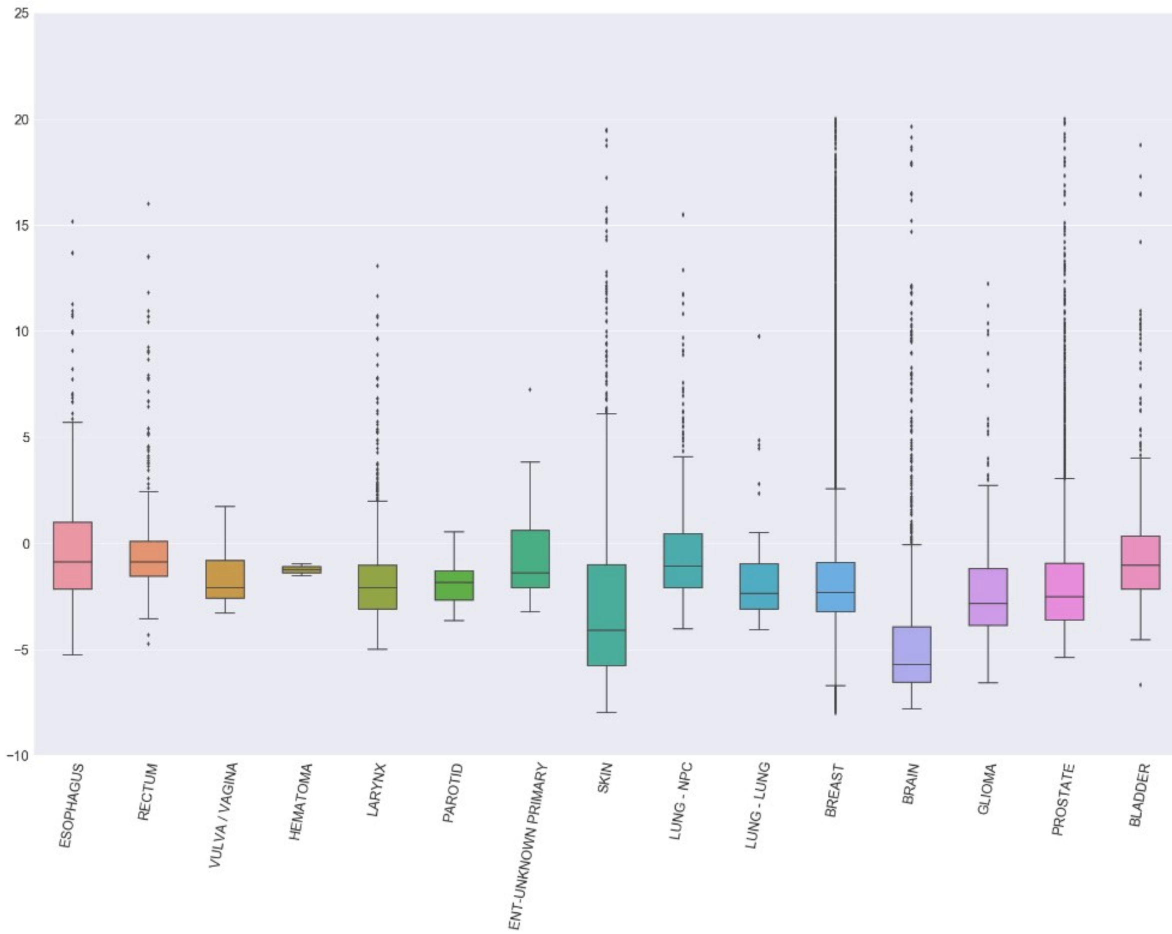

Figure 4.16 Residuals grouped by cancer category for 10 min class.

4.5.6 Experiments and Results

To evaluate our approach, we reschedule the patients of September 2017. We retain the same treatment days for each patient. We allocate new appointment times, while ensuring that all

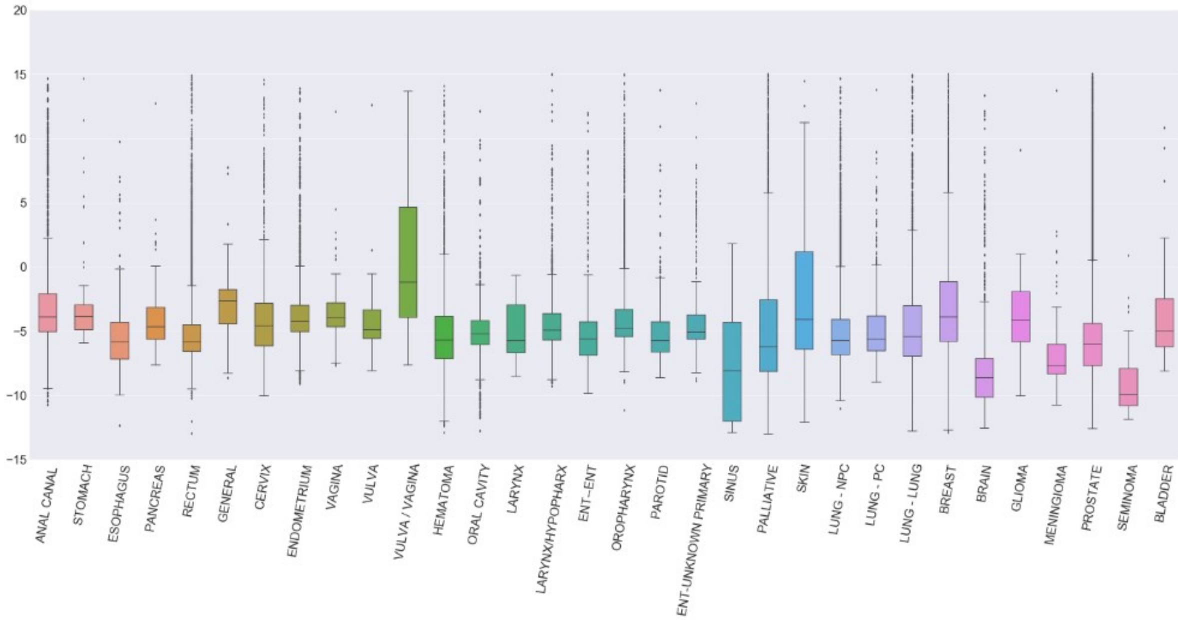

Figure 4.17 Residuals grouped by cancer category for 15 min class.

the treatments are done at the same time each day. To respect this constraint, we modify our prediction algorithm. We note that 20% of the patients change class over the course of the treatment. We assign them to their main class. We assume that the patient is ready when it is his turn to start the treatment. We perform 30 replications to simulate 14 scenarios with the two operational management rules.

Table 4.7 summarizes the results. The waiting-time columns give the means and standard deviations of the waiting times for the two rules. The first rule allows the patient to start the treatment early; however, the second rule forces the patient to wait until his appointment time. The overtime columns give the mean overtime for the two rules.

To assess our approach, we consider the first rule, which is currently applied at the CICL. The first grid gives the poorest results for the number of patients per day; although compared with the current CICL schedule there are three more patients per room. For some patient groups the length of the appointment interval is greater than the actual service time. For this reason, most patients are treated well in advance, and so the waiting time is zero. For the other three grids, ten more patients are treated per room. Since these grids have the same number of patients per day, we compare them in terms of waiting time and technologist overtime. There

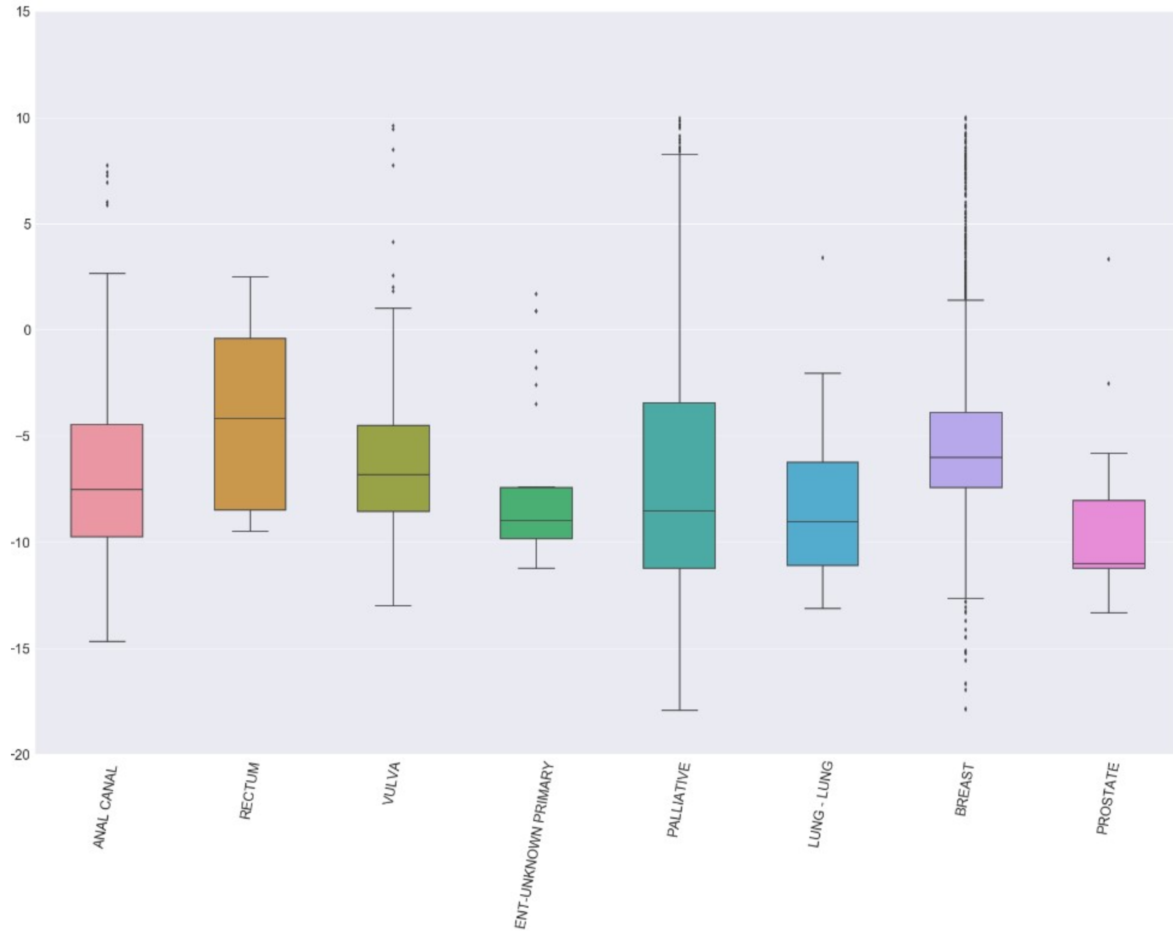

Figure 4.18 Residuals grouped by cancer category for 20 min class.

are only small differences between the three grids and between the sequencing rules.

For the waiting time, the SVF rule gives the best results for grids 2 and 4. For grid 3, the SMF rule is the best, and the alternation rules are the least efficient. For the overtime, the SMF rule is consistently the best. However, the average difference between the best rule in each case and the no-rule assignment is at most 0.15 min.

Grid 3 (alternates between 15 min green and red) outperforms the others, although the difference is small (0.4 min on average). Therefore, we recommend applying this schedule with the no-rule assignment.

Table 4.8 compares our solution with the current CICL schedule. In our solution, there is an increase of 10 patients per room per day, and a considerable improvement in the direct waiting

Table 4.7: Results of simulated scenarios

Scenario number of slots *[duration class]	Number of appointments per day	Sequencing rule	Waiting time 1st rule Mean	Waiting time 1st rule Std	Waiting time 2nd rule Mean	Waiting time 2nd rule Std	Overtime 1st rule	Overtime 2nd rule
12*[15]+22*[20]+1*[25]	140	SVF	0.231	1.438	0.751	2.506	0.016	0.166
		SMF	0.240	1.413	0.680	2.311	0.030	0.152
		No rule	0.224	1.416	0.680	2.346	0.013	0.114
12*[15]+29*[15]+1*[20]	168	SVF	2.137	5.587	4.549	7.705	2.512	5.118
		SMF	2.610	6.501	4.271	7.659	1.223	2.870
		No rule	2.423	5.818	4.630	7.828	1.506	3.047
12*[15,15]+17*[15]+1*[20]	168	SVF	1.827	4.814	3.635	6.511	2.221	4.543
		SMF	1.738	4.526	3.242	5.839	0.978	2.165
		No rule	1.857	4.669	3.481	6.228	1.188	2.581
		Variance Alternation	2.565	6.003	4.405	7.246	1.549	3.076
		Mean Alternation	2.151	5.345	3.508	6.362	1.650	3.066
Random sequence	168	SVF	2.102	5.217	3.892	6.837	2.287	5.279
		SMF	2.146	5.129	3.488	6.210	1.656	3.275
		No rule	2.154	5.119	3.566	6.384	1.661	3.242

time and the technologist overtime. There is a reduction of 4.5 min in the average waiting time as well as a remarkable decrease in the dispersion. Moreover, the average overtime is reduced by 6.6 min.

Table 4.8: Comparison of current and new schedule

	Number of appointments per day	Waiting time Mean	Waiting time Std	Overtime
Current schedule	128	6.45	7.35	7.87
New schedule	168	1.86	4.67	1.19

For the management rules, Table 4.7 indicates that it is always better to treat the patient as soon as the machine and technologists are ready. This decreases the waiting time and technologist overtime. For example, comparing grids 2 to 4, we see that the smallest difference between the two strategies is about 1.4 min for the waiting time and 1.2 min for the overtime; these values are seen with the SMF rule in grids 3 and 4 respectively.

4.5.7 Discussion

We have evaluated four schedules using five sequencing rules and two operational management rules. The sequencing rules have similar performance, so we recommend the simple no-rule assignment.

The best grid as measured by the performance indicators is the grid that alternates between the 10 min and 15 min classes, where the interpatient duration is added only to the first class. The alternation minimizes the variation in the waiting time; see Table 4.7. Our study demonstrates that allowing the patient to start the treatment early decreases the waiting time and technologist overtime.

Our schedule treats 10 more patients per day in each room and decreases the waiting time and technologist overtime. It will increase both patient and technologist satisfaction as well as the utilization of the center. In addition, it is easy to implement. The patient duration class is predicted using the CR tree, which is not difficult to apply. The grid has three duration classes, but since we add the interpatient duration only for the 10 min class, the grid has one 20 min slot, and the remaining slots are 15 min slots in two colors. Therefore, our schedule is characterized by flexibility and simplicity.

4.6 Conclusion

We have carried out a data-driven study to develop an efficient patient scheduling system. It increases the number of patients per day and decrease the direct waiting times and the overtime worked.

First, we developed a prediction model for the treatment duration. We applied several data mining and regression tools: general linear model, MARS, artificial neural networks, and the CR tree. The best model was provided by the CR tree, with an accuracy of 84%. The prediction model assigns a treatment time to each radiotherapy patient. Using the predicted durations, we designed new workday divisions and compared them using different patient sequencing rules. We found that the sequencing rules have only a small influence on the scheduling performance. The new schedule gives a 30% increase in the number of patients per day and a decrease in the waiting time and technologist overtime.

We conclude that the application of powerful tools such as data mining techniques can contribute to the design of a more efficient patient schedule. In addition, this study confirms that a nonblock scheduling system is realistic and effective, since the service time differs from one patient to another and depends on the treatment characteristics.

One of the limitations of our work results from the attributes used to predict the treatment duration. CICL has not provided patient details such as age and weight; these attributes can have a significant impact on the service time. In future work, we plan to include such features in the prediction model to improve its accuracy and consequently the effectiveness of the patient scheduling system.

CHAPITRE 5 ARTICLE 2: ON INTEGRATING PATIENT APPOINTMENT GRIDS AND TECHNOLOGIST SCHEDULES IN A RADIOLOGY CENTER

D. Ben Tayeb, N. Lahrichi et L-M. Rousseau

Cet article a été publié dans Health Care Management Science, pages 1-17
(21 Octobre 2022).

5.1 Abstract

Optimal patient appointment grid scheduling improves medical center performance and reduces pressure from excess demand. Appointment scheduling efficiency depends on resource management, and staff are a key resource. Personnel scheduling takes into account union rules, skills, contract types, training, leave, illness, etc. When combined with appointment scheduling constraints, the complexity of the problem increases. In this paper, we study the combination of the patient appointment grid and technologist scheduling. We present a well-detailed framework outlining our approach. We develop two versions of a mixed-integer programming model: integrated and sequential. In the first version, we elaborate the appointment grid and the technologist schedules simultaneously, while in the second version we generate them sequentially. We evaluate the proposed approach using real data from the MRI department of the Centre hospitalier de l'Université de Montréal (CHUM) radiology center. We study different scenarios by testing several technologist rules and planning construction methods. Obtained solutions are compared to the current CHUM scheduling approach.

Keywords : OR in health services, scheduling, patient appointment grid, radiology, policies.

5.2 Introduction

In recent years, increasing access requests and limited medical budgets have put pressure on hospitals to minimize costs while maintaining a high standard of care. Many healthcare facilities have costly equipment which is in high demand, especially in radiotherapy and imaging centers. A shortage of medical personnel results in machine underutilization, unmet demand for services and long patient waitlists. Between 1993-2003, the number of MRI exams and the number of computed tomography (CT) scans performed in Ontario, Canada increased threefold (Van Nynatten et Gershon (2017)). Canadians could wait an average of 4.8 weeks for a CT scan, and 9.3 weeks for an MRI scan (Bacchus et Mackenzie (2019)). Healthcare administrators are forced to manage their resources efficiently by reducing operational costs while ensuring patient satisfaction. The workforce represents an important source of direct costs. However, good management maximizes the number of patients seen per day and reduces new investments.

Schedulers in medical centers tend to adopt the simplest scheduling version that they can implement manually and easily. Appointment scheduling and staff scheduling are usually performed separately. Appointment grids are standardized and their elaboration is not contingent upon the human resource or material resource planning. As a consequence, machine schedules are static and the associated personnel planning is predetermined. Moreover, the number of personnel allocated to treat a given patient is usually fixed and depends on the shift type rather than patient status, especially in the case of technologists in the imaging and the radiotherapy centers.

In any healthcare center, the medical personnel have varying levels of training and skill, and their patients have different treatment characteristics. Some patient categories require the availability of appropriate personnel; for others, the presence of more than one caregiver or technician is needed to ensure a more efficient treatment. So while the elaboration of the appointment and staff schedules in a standardized manner is convenient, adding flexibility allows us to capture the complexity of the problem in real life, including the heterogeneity of patients and medical personnel.

Since one of the big challenges of healthcare research projects is their real-world application, we collaborated with the MRI department of the Centre hospitalier de l'Université de Montréal (CHUM) radiology center from the outset of our study, to determine the feasibility of our solutions. To ensure the inclusion of a wide range of real-world constraints, we consulted a multidisciplinary team of administrators, including: the head of the imaging department, a radiologist, administrative coordinators, as well as planners and assistants.

As with most medical centers, the design of the appointment grid and technologist schedules at the CHUM is performed manually and separately (Figure 5.1) by two different agents. The master appointment grid is pre-defined, while the technologist schedules are prepared for each period based on technologists' availability and skills. The technologist schedules are made first, then the functional grid is generated according to the modifications produced. Modifications related to the grid, such as machine shut down, lead to adjustment of the technologist schedules. Any change or adjustment in data requires communication between the schedulers. They then try to apply the fewest possible modifications, in an attempt to approximate the basic schedule.

In this paper, we propose a monthly (tactical-level) appointment grid and technologist scheduling approach, which takes demand into consideration. We provide an optimal allocation of personnel resources to maximize machine utilization and the number of patients seen per day. Our contribution is a managerial tool for healthcare administrators which guides their decision-making on staff structure, recruitment, working rules, and planning construction methods. Our research also helps managers compare their current process to more optimal processes, based on analysis of real data. This new information can subsequently be used to improve scheduling system efficiency. We define a well-detailed and structured framework composed of four steps:

1. Problem definition and formulation: We present two versions of an optimization model that considers the construction of the appointment grid and the technologist schedules: integrated and sequential. The integrated model consists of simultaneous appointment grid and technologist scheduling. In the sequential model, we elaborate technologist schedules first, then, design the appointment grid.

2. Data collection and investigation: The proposed approach is applied to real datasets from the MRI department in the CHUM radiology center. We use historical data spanning a period of three months. We lead a data-driven study of the impact of allocating a variable number of technologists to execute an exam category on the total number of patients seen per scheduling period.

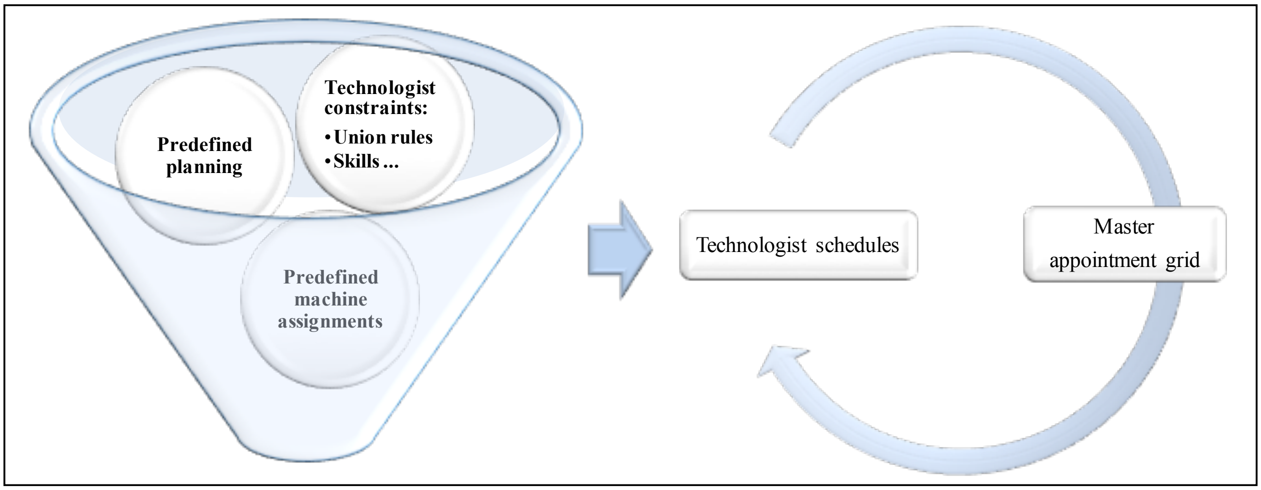

Figure 5.1 The CHUM scheduling approach

3. Experimental setting: We evaluate the two versions of the mathematical model while changing two CHUM standard scheduling elements: technologist working rules, and technologist planning construction methods. Our solutions are compared to the obtained solution by the CHUM approach.

4. Solution selection: We clarify the impact versus the challenges of applying each proposed scenario in real life, and we advise the CHUM administrators about which solutions could be implemented to improve efficiency.

The rest of the paper is organized as follows: Section 2 gives an overview of the relevant literature related to our work. Section 3 defines the problem statement and presents the proposed scheduling approach. Section 4 introduces the case study. In Section 5 we discuss the results. Section 6 concerns the solution selection plan. Finally, we conclude with a summary of the present study.

5.3 Literature review

We address two sides of the scheduling problem in this paper: appointment grid scheduling and technologist scheduling, which is a particular type of personnel scheduling.

The challenges associated with personnel scheduling depend on the level at which decisions are made. At the strategic level, we determine the size of the manpower, the skills, the contract, the number of trainees, etc. Tactical decisions include personnel schedules over some days or weeks, trainee planning and rotation, etc. However, at the operational level we present more detailed personnel schedules by attributing, for example, the tasks to perform and their sequence.

The personnel rostering or scheduling problem considers the tactical and operational levels of the planning horizon. This problem is divided into three different types: shift scheduling, days-off scheduling and tour scheduling, which integrates the first two classes. Ernst *et al.* (2004) define personnel scheduling as a well-defined process composed of a number of modules: demand modeling, days-off scheduling, shift scheduling, line of work construction, task assignment, and staff assignment. In the literature, authors apply different methods to model and solve the problem, such as mathematical programming approaches, constraint programming, decomposition, metaheuristics, simulation, etc. (Van den Bergh *et al.*, 2013; Ernst *et al.*, 2004; Alfares, 2004).

There are many challenges in the construction of personnel schedules, relating to the unique characteristics of the problem. An employer recruits its workforce with different types of contracts: full-time, part-time and casual. In some cases, the employees are from heterogenous sets and they do not have the same skills (Van den Bergh *et al.*, 2013), so specific tasks are assigned to the appropriate person. Moreover, there are several constraints to consider in modeling the problem, such as government regulations and employer rules.

In healthcare systems, the scheduling problem is most often studied in relation to nurse scheduling and physician scheduling; we refer the reader to Burke *et al.* (2004) and Erhard *et al.* (2018). The authors of these papers aim to effectively consider the complexity of the problem, taking into account fairness, preferences, and shift types. In fact, there are two main shift types: predefined and flexible. The first type has been studied widely in the literature with some authors studying overlapping shifts (Erhard *et al.*, 2018). With flexible shift types, it is possible to assign shifts with different start times and lengths. Brunner *et al.* (2009) introduce a physician scheduling model that combines shift scheduling, days-off scheduling and line of work construction, while permitting flexible shifts, breaks and overtime. In a subsequent paper they solve realistic instances by developing a branch-and-price algorithm (Brunner *et al.*, 2010). Stolletz et Brunner (2012) present a physician scheduling model with flexible shifts and fairness aspects. They reduce the set covering approach by generating all possible shifts in a preprocessing step using an algorithm. In the literature, the technologist scheduling problem has not attracted much attention, only a few published papers deal with this problem. Chen *et al.* (2016) propose a two-stage method for the allocation and scheduling of radiologic technologists. The first step determines the minimum required number of technologists; the second step establishes their schedules. Yuura *et al.* (2017) present a model for radiographer scheduling by integrating their skills and the trainee training. Vieira *et al.* (2018) optimize the allocation of radiotherapy technologists to multiple operations by considering stochastic patient arrivals. In these studies, the authors consider that the technologist coverage constraint to perform a procedure is fixed; they do not analyze the potential improvement by varying the number of staff for specific task or patient category.

Some studies combine the scheduling of resources with the scheduling of another component of the process, such as patients. Ogulata *et al.* (2008) study physiotherapist and patient scheduling over the workday. They maximize the number of treated patients while taking into account fairness among physiotherapists by balancing the workload between them. There are two studies concerning patient and resource scheduling in nuclear medicine. The procedures or tests in this speciality area of radiology are multi-stepped and require multiple human resources: technologists, nurses, physicians and managers. The first study (Pérez *et al.*, 2011) proposes an algorithm with a fixed procedure-resource assignment, and another algorithm that performs this assignment using an integer programming model for days with high demand. Pérez *et al.* (2013) extend the two algorithms to the stochastic online optimization, starting with the offline version.

For patient scheduling based on appointment systems, a redesign of the appointment grid according to patient categories, or service time types, leads to a more efficient scheduling system. Huang et Verduzco (2015) reclassify the patient scheduling groups and determine their time slot lengths. Van Sambeek *et al.* (2011) propose a new scheduling strategy to improve patient access time. They change the master appointment schedule by minimizing the number of block types, keeping only the important patient categories. Bentayeb *et al.* (2019) present a new appointment schedule based on a service time prediction model, which is elaborated using a data mining method.

In the existing literature, authors address only the patient appointment grid design, or they deal with patient scheduling with resource assignment at the operational level. They do not incorporate the tactical staff scheduling problem, taking into account the real constraints and personnel satisfaction. Moreover, the allocation of more than one technologist to treat a patient is not investigated. To the best of our knowledge, our study is the first to evaluate the performance of the system when the number of technologists assigned to a patient according to its category varies. In this paper, we aim to study the impact of the combination of the patient appointment grid and technologist scheduling on the number of treated patients. We evaluate different scenarios based on real data, by changing the technologist planning construction and the working rules.

5.4 Problem statement and modeling

In this section, we describe the characteristics of the appointment grid and technologist scheduling problem. We then present the mathematical formulation.

5.4.1 Problem statement

The appointment grid scheduling problem consists in determining the following for each exam machine: the type, the number and the sequence of the exams. In the technologist scheduling problem, we determine the following for each technologist: their days off, the daily machine assignment, and the start and break times during the workday. Technologist scheduling is complicated in real life. We are dealing with people, so we must consider their working conditions, contract types, leaves, illnesses, etc. Technologist scheduling interacts with patient appointment management. The integrated planning of these two elements represents the fundamental problem; however, the consideration of all of the real appointment and technologist constraints amplifies the problem's complexity.

In the radiology centers, specifically in the MRI department, there are different technologist groups classified by their work contracts. In a shift work system, the technologist can work in the morning, in the afternoon, or in the evening. Moreover, he can be on a full-time or a part-time contract that also determines the number of compulsory hours that he has to work every fortnight. This is variable in the case of part-time jobs; the number of hours may differ depending upon whether the technologist works six, seven or eight days per fortnight.

Technologists also differ according to their skills. The training process in the MRI department is long. The need to use more machines and the high volume of patients prevent technologists from being trained to do all exam categories.

Moreover, the MRI machines are not identical in terms of transmitted magnetic field, design, manufacturer, license, etc. Each machine is dedicated to performing specific sets of exam categories.

Assigning exam categories depends not only on technologist skills and machine types, but also on time periods. In fact, some exam categories require the availability of other medical personnel, so they must be executed on pre-determined days and during specific slots.

Furthermore, the number of allocated technologists influences the performance of the exam execution process. For some categories, assigning more than one technologist increases the number of patients treated per time period. For other categories, the assignment of multiple technologists is a requirement.

Most centers allocate technologists based on their total number per shift. The number of technologists allocated to a machine is uniform per shift. Assigning exam categories to machines is done independently by taking into account the opening and closing hours of rooms.

Considering all these constraints in an integrated way leads to many challenges. In fact, most medical centers including the CHUM radiology center perform scheduling manually, and schedulers prefer to keep a standard and predetermined scheduling format to minimize changes.

In this article, we propose a mathematical model which includes simultaneous appointment grid and technologist scheduling. We evaluate the proposed model using real data from the MRI department of the CHUM radiology center.

5.4.2 Model

We present a mixed-integer programming model to optimize the design of the appointment grid and the elaboration of the technologist schedules. We suggest two versions of the mathematical model: integrated and sequential. We first present the integrated model, followed by the sequential version.

We develop a model to schedule exam categories and technologists monthly. The planning horizon is 28 days ($j \in J$) over four weeks. Each week $i \in$ I consists of weekdays from Monday to Friday (we define W_i as the set of working days from Monday to Thursday), and weekend days $j \in J_W$. J_W=$J_{st} \cup J_{sn}$, where J_{st} is the set of Saturdays, and J_{sn} is the set of Sundays. The set of days J is divided into two subsets of 14 days: the first 14 days in the planning is J1, and the second 14 is J2.

Our model takes into consideration technologist working conditions according to union rules and contracts. Technologists have a determined number of consecutive days q that they can work. Technologist $t \in T_k$ cannot work more than k days over 14 days. In addition, there is a set of technologists $t \in T_W$ who cannot work on the weekend.

In our model, we use the terms "shift" and "planning". To avoid any confusion, we give their definitions: a shift is a period of time during the day or night when an employee or a group of employees is scheduled to work; however, a planning determines the start and end times of the shift, the slots for work and the break times.

We note F and P sets of shifts and planning respectively. For each shift $f \in F$, there is a set of associated planning $p \in P_f$. The present model formulation splits the day into slot $s \in S_d$ of length $d \in D$ in hours. To determine if a given planning p covers a slot s, we use a binary parameter l_{ps}.

The technologist assignment depends on his competence. Binary parameters b_{tc} and a_{tpm} define the possibility of assigning technologist t to a category c, or a technologist t to planning

p and a machine m respectively.

We assign the exam category c to a machine based on a binary parameter e_{cm}, and to a slot on a given day by a binary parameter g_{csj}.

The execution of each exam category needs m_c technologists. The number of technologists $h \in H$ that perform a category c determines the average number of executed exams of category c per hour n_{hc}. We limit the maximum number of technologists to assign by the maximum number of the set H.

In the radiology center, the demand is variable. Therefore, the model respects the required number of one-hour slots for each category (parameter r_c). Moreover, we consider the minimal coverage of active machines during a shift using a parameter o_f.

Our model decides:

- if category c is performed on machine m, slot s and day j, using binary variables x_{cs}^{mj};
- if planning p is assigned to technologist t and machine m on day j, using binary variables y_{pt}^{mj};
- day-off j of a technologist t, using binary variables γ_{tj};
- the number of technologists h who perform each category c, on machine m, slot s and day j, using binary variables β_{hcs}^{mj}; and

Table 5.1: Summary of parameters

a_{tpm}	equals 1 if planning p can be assigned to technologist t on machine m
b_{tc}	equals 1 if technologist t can be assigned to exam category c
e_{cm}	equals 1 if it is possible to assign category c to machine m
g_{csj}	equals 1 if it is possible to assign category c to slot s and day j
l_{ps}	equals 1 if planning p covers slot s
m_c	minimal required number of technologists to execute the exam category c
n_{hc}	average number of executed exams of category c per hour if performed by h technologists
o_f	minimal number of active machines in shift f
q	maximum number of consecutive days that a technologist can work
r_c	minimal required number of one-hour slots for each category c

- the number of treated patients on each slot s during day j on machine m, using variables z_s^{mj}.

The model considers the stability of:

- appointment grid, using binary variables δ_s^{mj}, equal to 1 if a change of category is made on slot s, day j, machine m;
- technologist schedule by machine, using binary variables α_{tj}^{m}, equal to 1 if technologist t changes a machine m on a day j; and
- technologist schedule by planning, using binary variables $\alpha_{tj}^{'p}$, equal to 1 if technologist t changes planning p on a day j.

Integrated model

Our objective function contains six sub-objectives. The first two maximize machine utilization and the number of patients scheduled during the planning period and the third sub-objective maximizes technologists' days-off during weekends. The fourth term in the objective function penalizes any change of category on the same machine during a day. The last two sub-objectives ensure the stability of the technologist schedules: we penalize any change of machine and planning for a technologist during a weekday.

We maximize

$$\begin{aligned}\sum_{s\in S}\sum_{m\in M}\sum_{j\in J}(\sum_{c\in C}x_{cs}^{mj}+z_s^{mj})+\sum_{t\in T}\sum_{j\in J_W}\gamma_{tj}-\sum_{s\in S}\sum_{m\in M}\sum_{j\in J}\delta_s^{mj}\\ -\sum_{t\in T}\sum_{j\in J}(\sum_{m\in M}\alpha_{tj}^{m}+\sum_{p\in P}\alpha_{tj}^{'p})\end{aligned} \tag{1}$$

Constraints (2) ensure that we can assign at most one exam category to a slot on a given machine during a day. Constraints (3) ensure that a technologist cannot work on more than one machine according to more than one planning per day.

$$\sum_{c\in C} x_{cs}^{mj} \leq 1, \quad \forall\, s \in S,\, m \in M,\, j \in J \tag{2}$$

$$\sum_{p\in P}\sum_{m\in M} y_{pt}^{mj} \leq 1, \quad \forall\, t \in T,\, j \in J \tag{3}$$

In order to respond to realistic demand, there is a minimum coverage to consider. Constraints (4) define the minimal number of each category over the scheduling period. Moreover, the imaging center determines the minimum number of active resources in each department on weekends, to cover planned appointments as well as emergencies. This is enforced by constraints (5).

$$\sum_{d\in D}\sum_{s\in S_d}\sum_{m\in M}\sum_{j\in J} d x_{cs}^{mj} \geq r_c, \quad \forall\, c \in C \tag{4}$$

$$\sum_{p\in P_f}\sum_{t\in T}\sum_{m\in M} y_{pt}^{mj} \geq o_f, \quad \forall\, j \in J_W,\, f \in F \tag{5}$$

In the MRI department, there are different exam categories that may require specific material and personnel resources. Before assigning an exam category to a slot on a machine, we have to secure the availability of the minimal number of appropriate technologists. This is ensured by constraints (6). Furthermore, we may execute an exam category only on the proper machines and during suitable slots and days. These two restrictions are enforced by constraints (7) and (8), respectively.

$$\sum_{p\in P}\sum_{t\in T} b_{tc} l_{ps} y_{pt}^{mj} \geq m_c x_{cs}^{mj}, \quad \forall\, c \in C,\, s \in S,\, m \in M,\, j \in J \tag{6}$$

$$x_{cs}^{mj}(1 - e_{cm}) = 0, \quad \forall\, c \in C,\, s \in S,\, m \in M,\, j \in J \tag{7}$$

$$x_{cs}^{mj}(1 - g_{csj}) = 0, \quad \forall\, c \in C,\, s \in S,\, m \in M,\, j \in J \tag{8}$$

Technologists are not all trained to work on every machine, nor do they all have the same work contract. Some technologists can work only the morning planning, while others work only the evening planning. Constraints (9) attribute the appropriate planning and machine to each technologist.

$$y_{pt}^{mj}(1 - a_{tpm}) = 0, \quad \forall\, p \in P,\, t \in T,\, m \in M,\, j \in J \tag{9}$$

Our model takes into account hospital and government regulations pertaining to technologists' work. Technologists cannot work more than five consecutive days ($q = 5$), nor can they work more than one weekend every two weeks. This is guaranteed by constraints (10) and (11), respectively. Constraints (12) ensure that the technologist on duty has to work both weekend days with the same assignment of machine and planning. Constraints (13) and (14) define the number of days that a technologist has to work over the scheduling period, according to his contract.

$$\sum_{p\in P}\sum_{m\in M}\sum_{j'=j}^{j+q} y_{pt}^{mj'} \leq q, \quad \forall\, t \in T,\, j \in \{1, .., max(J) - q\} \tag{10}$$

$$\sum_{p\in P}\sum_{m\in M} y_{pt}^{mj} + \sum_{p\in P}\sum_{m\in M} y_{pt}^{m(j+6)} \leq 1, \quad \forall\, t \in T,\, j \in J_{Sn} \tag{11}$$

$$y_{pt}^{mj} = y_{pt}^{m(j+1)}, \quad \forall\, p \in P,\, t \in T,\, m \in M,\, j \in J_{St} \tag{12}$$

$$\sum_{p\in P}\sum_{m\in M}\sum_{j\in J_1} y_{pt}^{mj} = k, \quad \forall\, t \in T_k \tag{13}$$

$$\sum_{p\in P}\sum_{m\in M}\sum_{j\in J_2} y_{pt}^{mj} = k, \quad \forall\, t \in T_k \tag{14}$$

Constraints (15) define binary variables γ_{tj}. In some cases, technologists cannot work on weekends because they are not yet trained, for example. Constraints (16) avoid assigning this group of technologists to the weekends.

$$1 - \sum_{p\in P}\sum_{m\in M} y_{pt}^{mj} = \gamma_{tj}, \quad \forall\, t \in T,\, j \in J \tag{15}$$

$$\gamma_{tj} = 1, \quad \forall\, t \in T_W,\, j \in J_W \tag{16}$$

One of the main objectives of this paper is to study the impact of the number of technologists allocated to perform an exam category on the number of patients seen. Constraints (17)-(19) are imposed to determine the number of patients treated in a given slot based on the number of assigned technologists.

$$\sum_{p \in P} \sum_{t \in T} l_{ps} y_{pt}^{mj} = \sum_{c \in C} \sum_{h \in H} h \beta_{hcs}^{mj}, \quad \forall\, h \in H,\, s \in S,\, m \in M,\, j \in J \tag{17}$$

$$\sum_{h \in H} \beta_{hcs}^{mj} = x_{cs}^{mj}, \quad \forall\, c \in C,\, s \in S,\, m \in M,\, j \in J \tag{18}$$

$$z_s^{mj} \leq \sum_{h \in H} \sum_{c \in C} dn_{hc} \beta_{hcs}^{mj}, \quad \forall\, c \in C,\, s \in S_d,\, m \in M,\, j \in J,\, d \in D \tag{19}$$

To enhance the quality of the proposed solution, we have to consider the stability of schedules. To track machine or planning changes for technologists during weekdays, we use variables α_{tj}^{m} and $\alpha_{tj}^{'p}$ in constraints (20) and (21), that are penalized in the objective function. On the other hand, the variables δ_s^{mj} in constraints (22) calculate the change of category from one slot to the next on the same machine during a day, which is minimized in the objective function.

$$\sum_{p \in P} y_{pt}^{mj} \leq \sum_{p \in P} y_{pt}^{m(j+1)} + \gamma_{t(j+1)} + \alpha_{t(j+1)}^{m}, \quad \forall\, t \in T,\, m \in M,\, j \in W_i, i \in I \tag{20}$$

$$\sum_{m \in M} y_{pt}^{mj} \leq \sum_{m \in M} y_{pt}^{m(j+1)} + \gamma_{t(j+1)} + \alpha_{t(j+1)}^{'p}, \quad \forall\, t \in T,\, p \in P,\, j \in W_i, i \in I \tag{21}$$

$$x_{cs}^{mj} + x_{c'(s+1)}^{mj} \geq \delta_{s+1}^{mj} + 1, \quad \forall\, c \in C,\, c' \in C,\, c \neq c',\, s \in S, \quad m \in M,\, j \in J \tag{22}$$

Constraints (23)-(29) serve to define the value ranges of our variables.

$$x_{cs}^{mj} \in \{0,1\}, \quad \forall\, c \in C,\, s \in S,\, m \in M,\, j \in J \tag{23}$$

$$z_s^{mj} \in \mathbb{N},\, \delta_s^{mj} \in \{0,1\}, \quad \forall\, s \in S,\, m \in M,\, j \in J \tag{24}$$

$$\beta_{hcs}^{mj} \in \{0,1\}, \quad \forall\, h \in H,\, c \in C,\, s \in S,\, m \in M,\, j \in J \tag{25}$$

$$y_{pt}^{mj} \in \{0,1\}, \quad \forall\, p \in P,\, t \in T,\, m \in M,\, j \in J \tag{26}$$

$$\alpha_{tj}^{m} \in \{0,1\}, \quad \forall\, t \in T,\, j \in J,\, m \in M \tag{27}$$

$$\alpha_{tj}^{'p} \in \{0,1\}, \quad \forall\, t \in T,\, j \in J,\, p \in P \tag{28}$$

$$\gamma_{tj} \in \{0,1\}, \quad \forall\, t \in T,\, j \in J, \tag{29}$$

Sequential model

The basic formulation of the sequential model is very similar to that of the integrated model, but we separate it into two models. The first one (the technologist scheduling model) considers the technologist schedule constraints. The second one (the appointment grid scheduling model) takes into account the exam category assignation constraints. In order to link the two models, and to propose a feasible and realistic solution, we add some notations and constraints, and we change the objective function. The complementary mathematical formulation is as follows:

We divide the objective function of the integrated model into two separate objective functions for technologist (1a) and appointment grid (1b) scheduling. We add a new term in the objective function (1a), that maximizes machine utilization by increasing the number of active slots.

- Technologist scheduling model

  We maximize

$$\sum_{s\in S}\sum_{m\in M}\sum_{j\in J}u_s^{mj}+\sum_{t\in T}\sum_{j\in J_W}\gamma_{tj}-\sum_{t\in T}\sum_{j\in J}(\sum_{m\in M}\alpha_{tj}^{m}+\sum_{p\in P}\alpha_{tj}^{'p}) \tag{1a}$$

  We consider the technologist scheduling constraints of the integrated model ((3), (5), (9)-(16), (20), (21), (26)-(29)), plus the following constraints:

$$\sum_{p\in P}\sum_{t\in T}l_{ps}y_{pt}^{mj}\geq u_s^{mj},\quad \forall s\in S, m\in M, j\in J \tag{2a}$$

$$\sum_{p\in P}\sum_{t\in T}l_{ps}y_{pt}^{mj}\leq max(H),\quad \forall s\in S, m\in M, j\in J \tag{3a}$$

$$\sum_{p\in P}\sum_{t\in T}l_{ps}y_{pt}^{mj}\geq hv_{sh}^{mj},\quad \forall h\geq 2, s\in S, m\in M, j\in J \tag{4a}$$

$$\sum_{d\in D}\sum_{s\in S}\sum_{m\in M}\sum_{j\in J}dv_{sh}^{mj}\geq r_h^{'},\quad \forall h\geq 2 \tag{5a}$$

  Constraints (2a) ensure that a slot on a given machine is active if at least one technologist is assigned to this time period. Constraints (3a) avoid assigning more technologists than allowed. Constraints (4a) and (5a) reserve capacity on appropriate resources to

Table 5.2: Summary of parameters and variables

	Parameters
$r_h^{'}$	minimal number of slots where it is possible to treat categories that require at least h technologists
Y_{pt}^{mj}	binary parameter represents the obtained value of y_{pt}^{mj} as output of the technologist scheduling model
	Variables
u_s^{mj}	binary variable, equal to 1 if at least one technologist is active on machine m, slot s, day j
v_{sh}^{mj}	binary variable, equal to 1 if at least h technologists are active on machine m, slot s, day j

perform categories that require more than one technologist.

- Appointment grid scheduling model

  We maximize

$$\sum_{s\in S}\sum_{m\in M}\sum_{j\in J}(\sum_{c\in C} x_{cs}^{mj} + z_s^{mj}) - \sum_{s\in S}\sum_{m\in M}\sum_{j\in J}\delta_s^{mj} \tag{1b}$$

  We consider the appointment grid scheduling constraints of the integrated model ((2), (4), (6)-(8), (17)-(19), (22)-(25)) including two modified constraints. The technologist assignations, which are the output of the technologist scheduling model, represent an input for the appointment grid scheduling model. We replace the variables y_{pt}^{mj} with the parameters Y_{pt}^{mj} in constraints (17) and (6), resulting in constraints (2b) and (3b).

$$\sum_{p\in P}\sum_{t\in T} b_{tc} l_{ps} Y_{pt}^{mj} \geq m_c x_{cs}^{mj}, \quad \forall c \in C, s \in S, m \in M, j \in J \tag{2b}$$

$$\sum_{p\in P}\sum_{t\in T} l_{ps} Y_{pt}^{mj} = \sum_{c\in C}\sum_{h\in H} h\beta_{hcs}^{mj}, \quad \forall h \in H, s \in S, m \in M, j \in J \tag{3b}$$

We have presented two versions of our appointment grid and technologist scheduling approach, in order to compare the proposed integrated version regarding the sequential method. The two versions are evaluated through different scenarios based on a case study.

5.5 Case study

The proposed scheduling models are applied using real data from the MRI department of the CHUM hospital, and are compared to the real schedules. Our data is from January, February and March of 2019.

5.5.1 Data collection

In order to provide input data for our scheduling model, we collect data related to exam category distribution, machine capability, exam assignment, technologist shift structure, number of available technologists, technologist skills, technologist posts, minimal coverage, etc.

In the hospital, two schedulers determine the appointment grid and the technologist schedules separately. The design of the grid is basically fixed. The number of each exam type is based on the category distribution according to the patient waiting list. The assignment and sequence of the exams are related to different constraints, including the exam categories and the machine capability. Since the technologist scheduling deals with humans, who can be unpredictable, it undergoes more changes and instability. From month to month, and sometimes from week to week, the number of technologists varies as a function of leave, training, recruitment, etc.

Figure 5.2 shows the MRI category distribution for the year 2018. There are seven main categories : neuroradiology, abdominal, musculoskeletal, breast, cardiac, breast biopsy and vascular. 20% of the department's capacity is reserved for research and emergency exams, which can fall under any MRI category. The neuroradiology and the abdominal exams account for more than 50% of the demand. The musculoskeletal and breast represent 13% and 7% of exams, respectively. Cardiac, breast biopsy and vascular are the minority exams, constituting less than 5% of MRI requests.

The CHUM center contains six MRI machines from three different manufacturers. Three machines produce a magnetic field of 3 Tesla, while the magnetic field generated by the other three is 1.5 Tesla. Assigning an MRI exam to a machine depends on its category. Neuroradiology and musculoskeletal exams may be performed on five out of six machines .

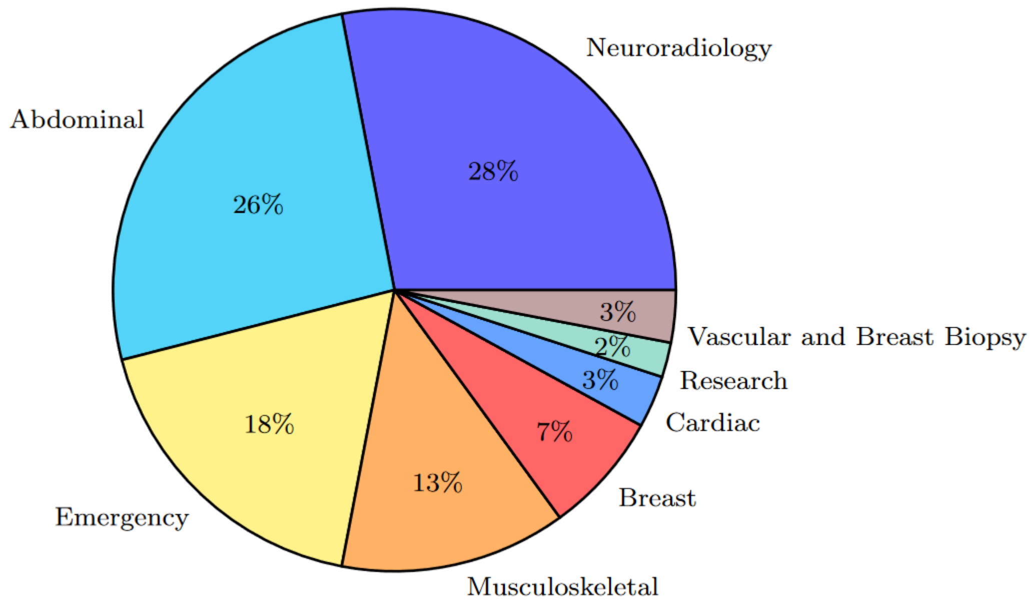

Figure 5.2 MRI Category distribution in the CHUM radiology center during 2018

Breast, cardiac and breast biopsy exams can only be executed on one machine. Abdominal MRI is performed on three machines, while vascular is performed on two machines.

There are some exams that can be done only on specific days or during specific time slots. Breast biopsy and cardiac exams require the availability of appropriate physicians. The breast biopsy MRI can only be performed on Tuesdays and Thursdays from 12:30pm to 2:30pm. The cardiac MRI can be performed on weekdays from 8am to 4pm. The breast exam requires an injection that may cause an allergic reaction in some patients. Exams requiring an injection cannot be performed after 10pm, to ensure that the center is still open and a technologist is present to take care of the patient in case of emergency.

Technologists in the MRI department work under rotating shifts. Shifts are successive, covering 16.5 hours per day, from 7am to 11:30pm. The morning shift starts between 7am and 9:30am, plus another shift that starts at 12pm. The start time for the evening shift is between 3pm and 4:30pm. According to the technologist contracts, the number of technologists for the morning shift is greater than the number of technologists for the evening shift. The scheduler generally assigns two technologists per machine during the morning shifts, and one technologist per machine during the evening. This can lead to underutilization of MRI machines in the evening.

Each technologist has three breaks per day: one break of one hour, and two breaks of 15 min. Technologists prefer to merge the two breaks of 15 min into one break of 30 min, and they take it at the end of their work day. Technologists who work the morning shift starting between 7am and 9:30am have one hour for their lunch break, and they leave 30 min before the end of their shift. However, the meal break in the remaining shifts is only 30 min, and the technologists' working day ends one hour in advance.

The opening and the closing hours are basically fixed for each machine. They are based on the predetermined technologist planning assigned to machines. For example, machines A and D are usually active only during the morning shifts. However, other machines are active in the morning and the evening shifts.

During the period of the study, the center had nineteen full-time technologists, and six part-time technologists that work only six, seven or eight days per two weeks. Technologists take their annual leave at different times during the year. Moreover, maternity and deferred salary leave leads to a technologist shortage during some periods. To cover the resource gap, some technologists accept shift changes or changes to their post type for a duration.

According to labor regulations, technologists can work one weekend over a two-week period. However, the center allows just one active machine for each weekend shift. In order to enhance patient access during the weekend, the availability of at least two active machines and two technologists in the morning shift has been required since February 2019. Moreover, the center tries to promote ongoing technologist training. Currently, all technologists can work on machines A, B and E, and they will be trained in the future to work on other machines and to perform complicated exams, such as cardiac exams.

5.5.2 Data investigation

In this case study, we aim to evaluate the impact of the number of technologists that perform MRI exams on service time. Figure 5.3 describes the patient exam execution process in the MRI department of the CHUM radiology center. Upon arrival to the center, the patient is registered. This appears directly in the CHUM appointment system, informing technologists that the patient is present. The patient changes his clothes and sits in the waiting room. During this period, the patient fills out questionnaires.

If the MRI machine is free, a technologist has to clean and prepare the room for the planned exam. The technologist locates the patient in the waiting room, and they head toward the preparation room. The technologist checks and fills in missing information in the patient questionnaires. He describes the MRI exam procedure and its duration. Some MRI exams require a specific injection. In particular exam categories, most patients are injected. In such cases the technologist has to prepare the patient for the injection by placing the catheter.

Once the patient is ready, the technologist guides him to the exam room. He positions the patient on the MRI machine. Then, the exam is performed. Once the exam is completed, the patient leaves the room.

In the MRI department, all exam categories can be executed by one technologist, except the breast biopsy exam, which requires at least two technologists. The availability of more than one technologist for some processes promotes system efficiency. For example, the preparation of the patient takes between 5 and 15 minutes. Sometimes the technologist has to fill out the questionnaire with the patient because he is unable to do it by himself or he doesn't have enough time. This increases the patient preparation duration. If two technologists were available for each exam, one could prepare and assist the patient with paperwork while the other one is monitoring the current patient in the exam room. Other tasks such as room preparation and patient positioning are likewise faster in the presence of two technologists.

Due to the limited number of technologists in the CHUM radiology center, it is not possible to assign two technologists to every MRI exam. In this study, we select the exam categories for which the allocation of more than one technologist has the greatest impact on increasing patient access.

In the MRI database, we find all the information related to each performed exam. However, only the technologist who executes the exam is registered. Patient appointment grid and technologist schedules are elaborated in two separate files by two different planners, making it is difficult to identify every technologist who attended a given exam. We run a statistical analysis on real data from January, February and March of 2019. Based on planning of

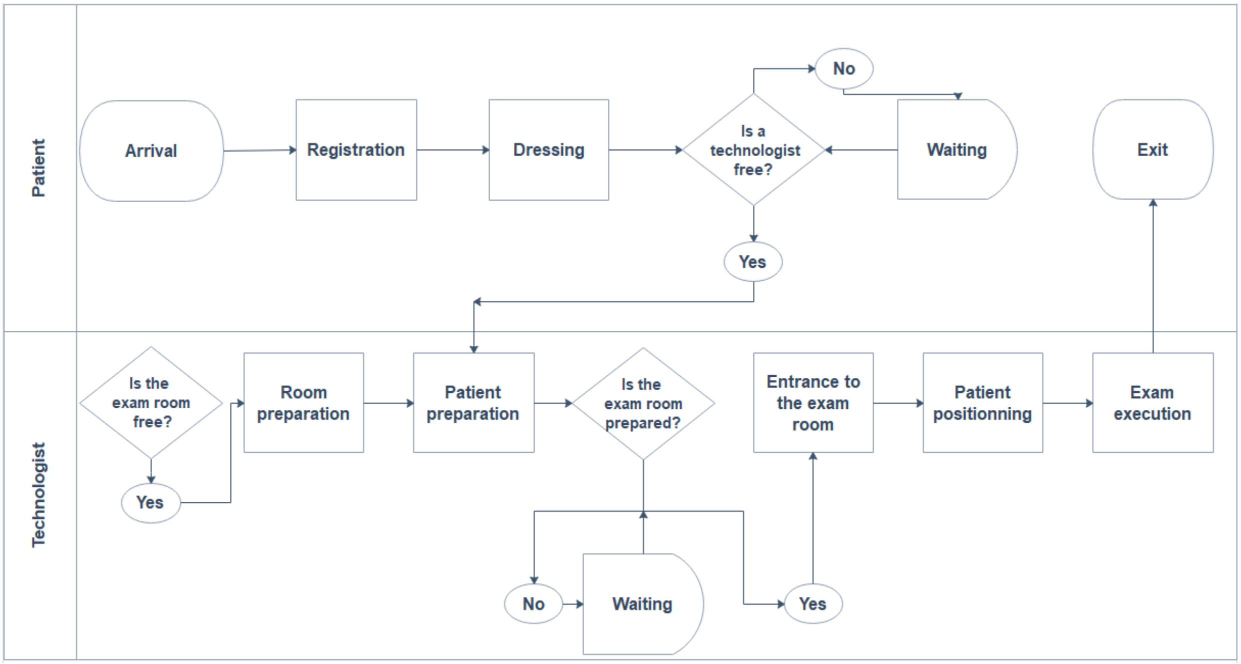

Figure 5.3 Patient exam execution process

assigned technologists to each machine, we determine the number of technologists for each exam, assuming that the break times corresponding to planning are respected.

Table 5.3 shows the average number of treated patients per hour for each category as a function of the number of allocated technologists. The difference is significant for some MRI categories, including neuroradiology, abdominal and breast exams; however, it is not the case for the remaining categories. This analysis helps us construct a more efficient appointment grid by assigning each exam to the convenient time period.

Table 5.3: The average number of treated patients per hour

MRI category	One technologist	two technologists
Neuroradiology	1.54	1.8
Abdominal	1.16	1.4
Musculoskeletal	1.46	1.48
Breast	1.48	1.9
Cardiac	0.73	0.87
Vascular	1.6	1.67
Breast biopsy	-	0.8

We use collected and prepared data to evaluate the current CHUM approach versus the performed scenarios.

5.5.3 Current approach

In the CHUM MRI department, planners elaborate technologist schedules based on predefined planning and machine assignments. Then, they check the feasibility of the appointment grid (Figure 5.1).

We generate technologist schedules using our sequential model with the current rules (CHUM-cur), allowing for each machine the specific planning determined by the CHUM. Following the current approach, we rectify the master grid according to the technologist schedules. For example, if we don't have an available technologist in a given slot or day, we cannot perform an exam category during this period.

5.5.4 Experimental design

This section concerns the experimental design. Our objective is to provide a managerial tool that evaluates the impact of administrator decisions on system performance. In order to determine which elements of the scheduling system to focus on in our study, we have analyzed the working conditions and rules at the CHUM, and spoke with managers about the challenges they face, the flexibility they have, and the framework limits. This discussion leads to three main areas where we saw potential for improvement in system performance: technologist weekend assignment frequency, technologist training and the planning construction method.

Technologist weekend assignment frequency According to union rules, technologists can work one weekend every two weeks. However, in practice, during the weekend, the CHUM opens only one or two machines during the morning shift, and one machine during the evening shift. Therefore, technologists work at most one weekend morning shift every seven or eight weeks, and one weekend evening shift every four weeks.

Technologist training In 2017, a new CHUM megahospital complex was opened, and personnel and equipment from three different hospitals (Hôtel-Dieu, Notre-Dame, and Saint-Luc) began to be reorganized and redistributed. Technologists were not trained to work on all machines. Although the center tries to train all technologists, the training process is long and the high demand does not allow launching several training sessions at the same time. The center prioritizes the increase of machine utilization by assigning the technologists according to their current skills.

The planning construction method The CHUM assigns specific planning for each machine. This facilitates manual schedule elaboration, and minimizes changes to the standard appointments grid. The master grid contains exam category blocs of 30 min. The two most important aspects of the planning construction method are the start and the break times.

In this study, we modify these three separate elements of the scheduling process, and we measure the impact on system performance.

We compare three technologist weekend assignment frequencies: the weekend assignment applied by the CHUM (CHUM_W), one weekend over two weeks (1/2), and one weekend over four weeks (1/4). We consider two cases of technologist training: technologists have current skills (TRAIN_{CHUM}), and all technologists are trained (TRAIN_{all}). We define a rule as a combination between the technologist weekend assignment frequency and the technologist training case; it represents the technologist working conditions in the center. Table 5.4 summarizes the six rules that we consider.

Table 5.4: Rules for technologist assignment

	Weekend frequency			Training	
	CHUM_W	1/2	1/4	TRAIN_{CHUM}	TRAIN_{all}
R0	✓			✓	
R1		✓		✓	
R2			✓	✓	
R3	✓				✓
R4		✓			✓
R5			✓		✓

Regarding the planning construction method, we introduce slots of 1h and slots of 30 min. We consider four planning construction methods. The idea is to allow more choices in the planning construction. We act on two factors: technologist start time and the corresponding meal time, while respecting the time interval constraints. We present two start time alternatives based on the offset between the possible planning of each shift: 1 hour, or 30 minutes. The offset is defined as the interval of time between two successive shift start times, it shows how start times of technologists are spaced. Let D_{meal} be the duration between the start time and the meal break. We consider two cases depending on the variation of D_{meal}. In the first case (D_{meal} fixed), D_{meal} is equal to four hours in the morning shift, and two hours in the evening shift. In the second case (D_{meal} variable), the values of D_{meal} in the morning shift vary between two hours and six hours; in the evening shift they vary between two hours and three and a half hours.

In Table 5.5, we show the four possible planning construction methods based on the changes to technologist start and meal times. P1 can be seen as the less flexible planning where the D_{meal} is fixed and the start time offset equals one hour. On the contrary, P4 represents the planning construction method with the highest flexibility; We allow variable duration between the start time and the meal time, and start times are spaced by 30 minutes.

In this paper, we compare the performance of the sequential and the integrated models for each scenario. A scenario is defined as a combination of a rule and a planning construction method (R_iP_j).

Table 5.5: Planning construction methods

	Start time		Meal time	
	1h offset	30 min offset	D_{meal} fixed	D_{meal} variable
P1	✓		✓	
P2		✓	✓	
P3	✓			✓
P4		✓		✓

5.5.5 Performance indicators

The quality of the proposed solution is measured based on the number of patients, schedule stability, and machine utilization.

Number of patients The number of patients treated over the planning horizon. In the application of our model, we estimate this number based on the number of technologists allocated to each time slot, using Table 5.3. Since demand and wait times for appointments are high, we assume that we can fill all available slots.

Schedule stability Schedule stability is a factor of schedule quality. Maximizing the schedule stability involves minimizing changes to the schedule. For technologist scheduling, stability is determined by the number of changes made to planning or to machine assignment for a technologist from day to day, during weekdays. After a technologist day-off, we allow a change of machine or planning. In appointment grid scheduling, stability is determined by the total number of category changes from one slot to the next one, on the same machine, during a day.

Machine utilization The ratio between the time period when the machine is active and the machine capacity, during opening hours. A machine is active when a technologist is available to perform exams.

5.6 Results and discussion

In this section, we use a series of scenarios to evaluate the performance of the two versions of the proposed model, by combining possible planning and rules as described in Section 4.4. Each scenario is applied to three data sets from real data collected in the CHUM radiology center. The data are from January, February and March of 2019.

5.6.1 Results

The aim of this study is to compare the quality of the results obtained from the integrated and the sequential models, regarding the planning structure and the rules. We carry out 144 computational experiments applying the two versions of our model on three data sets, and all rules and planning construction methods defined in Tables 5.4 and 5.5. The models are coded with JuMP (Dunning *et al.*, 2017) for mathematical optimization in Julia. They are solved by CPLEX Solver, using a PC with an Intel Core i7 2.80 GHZ and 16 GB RAM processor. The maximum allowed computational time is eight hours. The average computational time varies between about one and a half hours and four and a half hours, depending on planning construction method. The increase in the number of choices when elaborating planning increases the computational time.

Figures 5.4, 5.5 and 5.6 represent the computational results. We evaluate the performed scenarios based on the performance indicators described in Section 4.5.

Number of patients Both versions of the model give good results in terms of the number of treated patients compared to the CHUM's current approach. There is a significant increase by applying rules R1 and R4, with a minimum value of 10%. The flexibility in the planning construction method leads to the best planning combination during workdays to increase the number of patients. P4 promotes patient access more than other planning construction methods. There isn't a big difference between methods P2 and P3.

Schedule stability The quality of the obtained solution in terms of stability varies by model as well as by indicator. Category changes increase when the sequential model is applied, reaching a maximum value of 156. In general, the more choices we add to planning construction, the more category changes we see in the grid. The difference between the

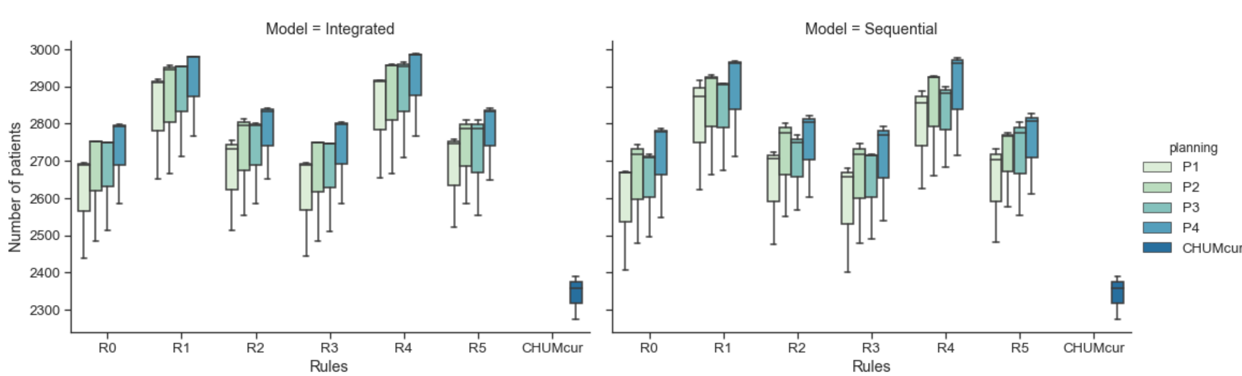

Figure 5.4 Number of patients

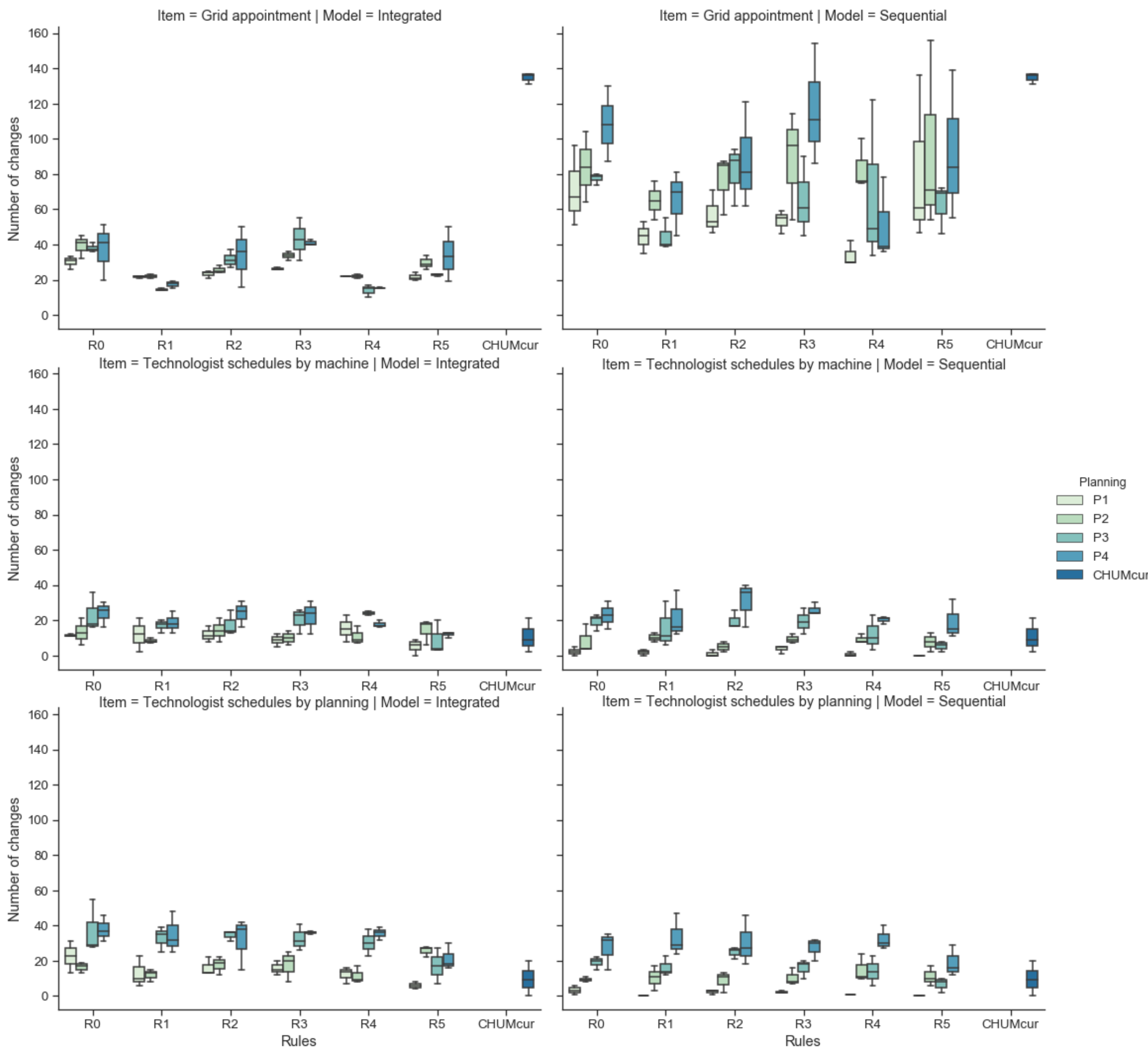

Figure 5.5 Schedule stability

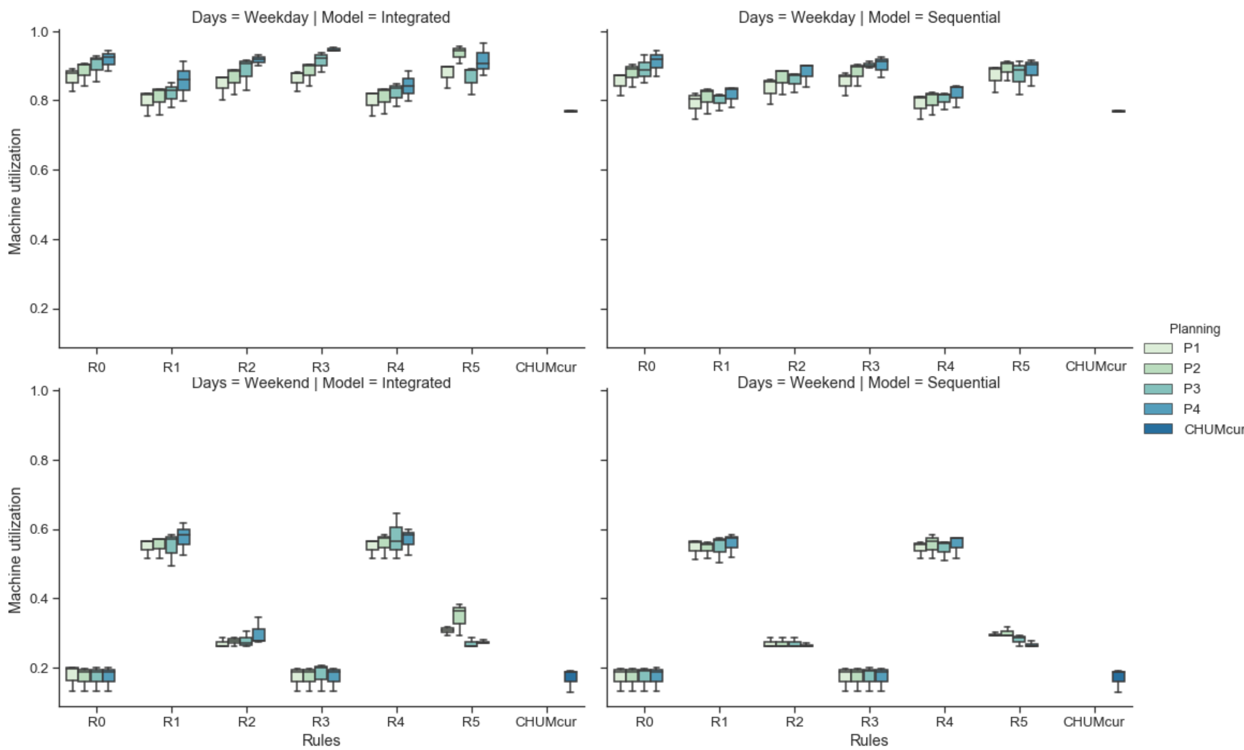

Figure 5.6 Machine utilization

average value of P1 and P4 is 30 for the sequential model. There is only a slight difference between the sequential and integrated models concerning the machine stability. However, the gap is larger for planning stability, with an average value equal to 10. The sequential model gives us the best results. In the sequential model, we start with the construction of the technologist schedule, with the objective of maximizing stability. Afterwards, we elaborate the appointment grid. We find that technologist schedule stability is much more important than the grid stability in terms of overall system performance.

Machine utilization Machine utilization depends on the planning construction method. The method that considers more choices in the planning of start and meal times increases machine utilization. It helps to find the best mix of planning that maximizes utilization. Rules that allocate more technologists during weekends (R1, R2, R4, R5) decrease the machine utilization on weekdays by about 8%. However, they increase utilization on weekends by 25%. Due to the limited personnel resources, there must be a balance between machine utilization on weekends versus weekdays. Rules R1 and R4 produce the maximum values for machine utilization during weekends, but reduce utilization during the weekdays.

All tested scenarios outperform the current CHUM procedure. The ability to modify the grid design, and to assign different planning and a variable number of technologists to each room, leads to improved performance of the technologist schedules and the appointment grid.

To present a fair evaluation of our approach, we measure the projected gain from implementing our proposed models, compared to the CHUMcur. We calculate the potential gain according to the best attained value. We observe an increase of 20% in terms of the number of treated patients, and a decrease of 92% for category changes, 20% for machine utilization improvement during the weekday, and 47% during the weekend. Concerning technologist schedule stability, we can produce schedules with zero machine or planning changes, especially in the sequential model.

5.6.2 Discussion and managerial insights

The experiment confirms that both the sequential and integrated models largely outperform the scheduling outcomes obtained by the CHUM approach.

As mentioned earlier, the CHUM approach fixes the planning assignment to each machine, and the number of allocated technologists to each shift. It also limits modifications according to the master grid. In contrast, our proposed model provides the ability to assign any planning to any room, and to generate the appropriate appointment grid depending on the technologist schedules. Our work showcases to hospital managers a better understanding of their data and how to utilize it to improve their scheduling system. We validate the direct impact of the number of allocated technologists to each exam. Including this information while preparing the scheduling presents an opportunity for performance improvement, even with the manual scheduling elaboration.

We show that adding more flexibility to technologist planning construction improves the performance of the generated schedules. The gap in the number of treated patients between the planning with the least flexibility (P1) and the planning with the greatest flexibility (P4) is approximately 4% on average. Moreover, rules influence the scheduling efficiency. There is a slight improvement in the number of treated patients by applying the rules that involve training all the technologists: R3, R4, R5 compared to rules R0, R1, R2 respectively. In fact, in CHUM, most technologists are already trained to perform the major categories. If the exam category distribution change, we could find a remarkable improvement in the case of trained technologists. Rules R1 and R4 represent the perfect case of the CHUM working conditions: Technologists can work one weekend over two weeks. The application of these rules increases the number of treated patients by 7% compared to R0, which represents the current CHUM rule.

In general, both the sequential and integrated models work well and produce good results. The sequential model has superior performance for some indicators such as technologist stability, while for other indicators, such as grid stability, the sequential model does not perform as well. The sequential model by its nature does not simultaneously consider all of the performance indicators. Once the technologist schedules are generated, maximizing the machine utilization and schedule stability, they remain unchanged. The number of treated patients is considered in the second step of the model, which is the generation of the grid. In order to maximize gain while respecting all exam category constraints, the sequential model is forced to make more category changes in the grid. However, the integrated model simultaneously combines the elaboration of technologist schedules and the appointment grid. Although this version of the model is harder to solve, it takes into consideration all performance indicators. The simultaneous model is designed to find a balanced trade-off between all of the sub-objectives, thereby increasing the quality of the obtained solution, as shown by performance indicators.

5.7 Solution selection

CHUM managers were interested and motivated to apply the present study to their work. We presented an evaluation of our integrated and sequential models through different scenarios that consider the proposed planning (P1, P2, P3, P4) and rules (R0, R1, R2, R3, R4, R5) applied to real data from the CHUM MRI department.

The results show that the integrated model leads to good overall performance in the elaboration of technologist schedules and the appointment grid. Our evaluation of this model demonstrated favorable results across all performance indicators. The CHUM managers decided to select the integrated model, and chose the planning structure and the rule based on the number of treated patients.

To assist the CHUM managers with implementing the model, we classified obtained solutions based on the effort required to implement them, and their impact on the center performance. We created an impact-effort matrix, which is a decision-making tool that utilizes graphic representation of a problem. The objective is to prioritize solutions with high impact and quick gain, to determine the most efficient applications.

The impact axis in Figure 5.7 represents the percentage gain in the number of treated patients compared to the CHUMcur. The effort axis represents the difficulty of implementing the solution in the CHUM center. The effort values were attributed during a meeting with the CHUM managers.

Table 5.6 summarizes the effort values. P1 is the easiest planning construction method to apply, the start time and the meal time are fixed, so the number of planning is reduced and the scheduler can attribute them readily. The CHUM managers consider that adding any degree of flexibility in the planning construction method has the same impact in the solution implementation, for this reason P2, P3 and P4 have the same effort value. Regarding the applied rules, increasing the technologist weekend assignment frequency and training all technologists increase the implementation difficulties.

Our work presents different possible options to enhance the center performance depending on their goals and work conditions. To avoid the challenges related to technologist training, using only R0, R1 and R2 give good results compared to the CHUM current rules. Moreover, changes in the planning construction methods in order to add the flexibility in the technologist start time and meal time lead to remarkable improvement. Based on our results, we also confirm that encouraging technologists to work more weekends increases the number of patients scheduled even without changing the current construction planning method. Although the theoretical number of worked weekends is one over two weeks, technologists currently work one weekend over eight weeks. Leadership will definitely be needed to implement this

Figure 5.7 Impact-effort matrix

Table 5.6: Effort values by rules and planning

Rules and planning	Effort values
R0	1
R1	8
R2	4
R3	6
R4	10
R5	9
P1	2
P2	3
P3	3
P4	3

effort. Note that we have evaluated the impact of one weekend over four weeks to help implementing the change gradually.

Figure 5.7 shows that solutions found by applying R0 with P1, P2 and P3, produce a low impact with low effort. This is logical, because R0 represent the CHUM current rule. Rules R3 and R5 combined with P1 and P2 is a waste of time when considering the low impact of that solution. The application of R1 and R4 is integrated among major projects that need complex activities. The establishment of R2 with P4 is considered a quick win project, with high impact from low effort.

The arrows on the grid indicate the hierarchical direction of solution recommendations. Firstly, we are interested in actions that ensure high impact and require low effort. We propose starting with the current CHUM rules and to only change the planning construction (R0P4). Next, we move to R2P4 by increasing the technologist weekend assignment frequency to one weekend over four weeks. The second step represents the application of projects that need more effort and have bigger implications for staff. We suggest increasing the technologist weekend assignment frequency (R1P4). On average, R4 does not clearly outperform R1; But the integration of technologist training (R4P4) increases slightly the gain and reinforces the center performance in case of exam category distribution changes or technologists shortage. The proposed solution implementation plan allows CHUM managers to improve their scheduling system efficiency and to enhance patient access to the center, while avoiding an abrupt transition for their staff.

5.8 Conclusion

In this paper, we introduce a structured framework to improve the scheduling system in a healthcare center. The proposed approach is defined step-by-step, starting with the problem definition and resolution, to the experimentation and the solution selection. We present the problem of technologist scheduling and appointment grid design. In most healthcare centers, these two elements are considered separately. However, in this paper, we study the simultaneous scheduling of technologists and the patient appointment grid. We take into account the real-life constraints of the problem. We develop a mixed-integer programming model with two versions: integrated and sequential. The integrated version combines the construction of the appointment grid and technologist schedules. In the sequential version, we start by scheduling the technologists, and subsequently elaborate the appointment grid. We examine the impact of the number of technologists allocated to execute the exam on the total number of treated patients. The aim of this study is to evaluate the two versions of the model, using real data from the MRI department of the CHUM radiology center. The evaluation is performed based on the current scheduling approach of the center, by testing several technologist working rules and planning construction methods. The solution performance is measured by the number of treated patients, schedule stability and machine utilization.

The two versions of the model give good results compared to the CHUM scheduling approach. However, the sequential model gives satisfying results according to some indicators, such as the technologist schedules stability, to the detriment of other indicators, such as the appointment grid stability. This is due to the non-integration of all objectives simultaneously. Added flexibility in the technologist planning construction, and an increase in the frequency of technologist weekend assignments improve the scheduling system performance.

In this article, we applied the proposed approach on a real case of an imaging center; in the future, we plan to extended it to other contexts where the patient scheduling is based on the appointment grid and where the number of allocated technologist to each patient influences service time.

CHAPITRE 6 ARTICLE 3: RADIOTHERAPY PATIENT RESCHEDULING UNDER MACHINE BREAKDOWN

D. Ben Tayeb, N. Lahrichi et L-M. Rousseau

Cet article a été soumis pour publication dans Health Care Management Science.

6.1 Abstract

Unexpected events in a dynamic healthcare environment can create disruption in patient scheduling. When patient treatment requires the use of specific machines, and those machines breakdown, the disruptions can be enormous. In this paper we consider such a situation with the case of radiotherapy, where patient treatment is based entirely on machine utilization. We present a rescheduling approach for radiotherapy patients. We define a general rescheduling framework that includes three rescheduling options (delay patients, overbook patients, use overtime), and allows variation of the priority of those options. We also generate several scenarios to evaluate our proposed rescheduling approach by varying the duration of machine breakdown, and the initial fill rate of the schedule. Our objective is to present a rescheduling decision tool that is adaptable to individual healthcare manager priorities, and expectations regarding performance.

6.2 Introduction

Radiotherapy is an important step in cancer patient treatment. It is a complex process that consists of two main phases: pretreatment and treatment. During the pretreatment phase, the treatment area is determined, and a patient treatment plan is created. Radiotherapy treatment is divided into several sessions delivered over a one day period using a specialized machine called a linac (linear particle accelerator), and faces many challenges. Each patient has preferences regarding their appointment time, but must respect their treatment plan, and the due date assigned by their priority classification. Additionally, oncology centers have requirements, such as the requirement that administration of all patient treatment sessions on the same machine during the same slot are done by the same technologist. These challenges are matched with other objectives, like the maximization of the machine utilization and patient treatment access. Efficient patient and resource management can reduce long patient waiting lists, and allow for better utilization of expensive equipment like radiotherapy machines.

There are similarities between the processes in manufacturing facilities, and the processes in medical facilities. Both aim to offer high quality services, satisfy their clients, and minimize their costs. In order for management to generate efficient plans, both facilities must optimize their resource utilization. However, unexpected events can disturb initial plans, and deteriorate the quality of pregenerated schedules. Ouelhadj et Petrovic (2009) classify such events into resource-related (machine breakdown, operator illness, etc.), and job-related (job cancellation, due date changes, etc.). In this step of the process, rescheduling is required to update the schedule due to the disruptions. Vieira *et al.* (2003) define rescheduling as the process of updating an existing production schedule in response to disruptions or other changes. In order to evaluate the impact of rescheduling and its efficacy, performance measures have to be considered. In the literature, performance measures are divided into three categories: measures of schedule efficiency, measures of schedule stability, and costs (Vieira *et al.*, 2003).

In healthcare centers, several rescheduling factors can appear, such as: machine breakdown, urgent patients, newly arrived patients, treatment cancellations, etc. In this article, we focus on the case of machine breakdown. In general for machine breakdown, administrators in medical centers dispatch the canceled patients without following any clear or predetermined procedure. They look for quick solutions instead of optimal solutions. Furthermore, this commonly results in a lack of documentation, no measurement of system performance after rescheduling, and no accounting of the resulting costs.

To the best of our knowledge, we are the first to study patient rescheduling due to machine breakdown in a radiotherapy center. We define three main patient rescheduling options: delay patients, overbook patients, and use overtime. Based on those three rescheduling decisions, we present a general rescheduling framework in terms of their level of priority. We propose a multi-objective approach that aims to minimize the number of delayed patients, patient waiting time, and use of overtime. We develop an optimization model, and a simple heuristic that is improved using local search operators. The proposed approach considers our three objectives in a hierarchical manner; we treat the objectives one by one consecutively based on a given rescheduling decision prioritization sequence. We also study how the priority sequence in a radiotherapy center changes when a machine breaks down. Our contribution is a decision-making tool that will help managers understand the impact of each rescheduling option on overall solution quality under different working conditions. We collaborate with the Centre Intégré de Cancérologie de Laval (CICL). To evaluate our proposed rescheduling framework, we simulate several scenarios by changing the duration of machine breakdown and the initial fill rate of the schedule. The results demonstrate that integrating the local search technique improves the overall solution quality obtained by applying the heuristic, and the optimization model validates that our heuristic gives generally acceptable results. This varies depending on the performance indicator being considered and its level of priority. Finally, we present a detailed analysis of the results to help managers understand the potential impact of each option available to them, safely plan for the expected duration of machine breakdowns, and select the best decision for maximizing the performance of their overall rescheduling strategy.

The rest of the paper is organized as follows: Section 2 defines the problem statement, and gives an overview of literature relevant to our work. Section 3 details the rescheduling approach. Section 4 introduces the experiments, and presents the results. Section 5 analyzes the results, and then we conclude with a summary and the limitations of this study.

6.3 Problem statement and literature review

Managers in radiotherapy face several challenges while planning patient scheduling. Each patient has several treatment sessions, each of which has to be administered on a specific machine (linac). Schedulers must respect the radiotherapy treatment and clinical constraints, while also considering optimal resource utilization and patient satisfaction.

Unexpected events like machine breakdown disturb not only the initial patient schedule, but also technologist planning. When a radiotherapy machine breaks down, schedulers have little time to consider possible solutions. In order to solve the problem as quickly as possible, they only consider patient treatment plan constraints, without considering clinical and technologists priorities. The intent of the schedulers is to maintain the stability of the initial schedule. To reschedule canceled patients, the schedulers allow overbooking, moreover, the radiotherapy centers not working three shifts per day use the maximum capacity buffer of overtime, if there are still patients not rescheduled yet, they can delay patients with low priority.

In the literature, rescheduling is widely studied, especially for manufacturing systems. We refer the reader to four main literature reviews (Uhlmann et Frazzon, 2018; Li et Ierapetritou, 2008; Ouelhadj et Petrovic, 2009; Vieira *et al.*, 2003). In a dynamic manufacturing environment, unexpected events can disrupt the production process. In order to minimize their impact on overall efficiency, changes in the production schedule need to be considered (Uhlmann et Frazzon, 2018). Examples of uncertainty in the production scheduling process include unexpected changes such as rush order arrivals, order cancellations, and machine breakdowns (Li et Ierapetritou, 2008). Ouelhadj et Petrovic (2009) classify dynamic scheduling into three categories: completely reactive scheduling, predictive-reactive scheduling, and robust proactive scheduling. In manufacturing systems, predictive-reactive scheduling is the most commonly used approach for handling events that disturb production. Predictive-reactive scheduling is a scheduling/rescheduling process.

Expectations for efficient scheduling in manufacturing systems don't differ in healthcare centers. Efficient scheduling must be able to handle dynamic changes in healthcare systems, such as patient lateness, resource unavailability, or equipment failure (Mageshwari et Kanaga, 2012). Rescheduling in healthcare has been widely studied, including personnel rescheduling and patient rescheduling. Clark *et al.* (2015) present a literature review for nurse shift rescheduling. They consider shift rescheduling to be more challenging than shift scheduling. Gross *et al.* (2018) is the first paper to study physician rescheduling; they use a mixed-integer linear program (MILP) to update duty and workstation rosters when there are personnel absences.

For patient rescheduling in chemotherapy, Gunasekaran *et al.* (2020) confirm that nonconformity to chemotherapy schedules is widespread in real life scenarios. This can be due to treatment complications, patients' social schedule conflicts, or facility administrative reasons. Their study analyzes the impact of modifications to chemotherapy schedules on breast cancer patients. 69% of patients included in this study experienced schedule modification, and those patients had a 2.34 times higher risk of death. Condotta et Shakhlevich (2014) propose a multilevel template for handling patient appointment scheduling for chemotherapy treatment. They model the problem using a MILP, with the goal of minimizing patient waiting times and balancing nurse workloads. Hooshangi-Tabrizi *et al.* (2020) consider an online scheduling problem for booking the chemotherapy patients' multiple requests. They propose an adaptive and flexible scheduling approach that treats appointment requests dynamically, and takes patient preferences into account. Condotta et Shakhlevich (2014) and Hooshangi-Tabrizi *et al.* (2020) use a daily rescheduling procedure to better allocate resources when unexpected events occur during a given day.

Radiotherapy patient scheduling has also been heavily studied. In order to solve radiotherapy scheduling problems, most authors have used mathematical programming techniques, while some used heuristics and metaheuristics (Vieira *et al.*, 2016). Conforti *et al.* (2008) propose an optimization model for a radiotherapy patient scheduling by accounting for their priorities in the waiting list. They maximize the number of patients being scheduled by rescheduling patients that have already started treatment. They extend the proposed model by taking into account patient availability (Conforti *et al.*, 2011). Conforti *et al.* (2010) adopt a non-block scheduling approach for radiotherapy treatments. They propose an integer linear program that improves linear accelerator utilization, and the number of scheduled patients. Legrain *et al.* (2015) develop an online stochastic optimization approach for radiotherapy patient scheduling. By solving real instances provided by the CICL, they demonstrated that their proposed solutions outperformed those of the cancer center. Vieira *et al.* (2020) consider patient time window preferences when scheduling appointments for radiotherapy treatment. They combine a MILP with a heuristic, and are able to solve intsances of the problem from large health centers in a reasonable amount of time. Petrovic *et al.* (2006) propose two constructive algorithms for radiotherapy treatment booking based on patient prioritization. The first algorithm books patients forward from the earliest possible start date, and the second algorithm books backward from the due date. Petrovic et Leite-Rocha (2008) present four constructive approaches for radiotherapy scheduling which they improve using an algorithm based on the metaheuristic GRASP. Kapamara *et al.* (2006) characterize radiotherapy patient scheduling as a job shop problem which is NP-hard; they propose several metaheuristic approaches for solving this type of problem. Bentayeb *et al.* (2019) develop a prediction model for radiotherapy treatment duration using data mining and regression methods. They propose a new design for appointment grids based on their model, and evaluate it using several patient management and sequencing rules. Saure *et al.* (2012) formulate a dynamic multi-appointment patient scheduling problem for radiotherapy treatment using a Markov Decision Process. They employ linear programming to approximate dynamic programming. They propose policies for efficient treatment capacity allocation and wait time reduction. Gocgun (2018) proposes a model similar to Saure *et al.* (2012), but he also considers cancellation of treatments, and uses a simulation-based approximate dynamic programming approach to solve the model.

In this paper we are the first to study radiotherapy patient rescheduling problem due to machine breakdown. We define the main rescheduling decisions adopted in cancer centers. We evaluate the priority sequence of rescheduling decisions by applying a simple heuristic and an optimization model.

6.4 Methodology

In order to reschedule the patients in a radiotherapy center when there is machine breakdown, we apply an optimization model and heuristic that follow a rescheduling framework presented in this section. Our approach proceeds by updating the initial schedule of the radiotherapy treatment to handle machine breakdown.

6.4.1 Rescheduling framework

In this article, we present a managerial decision tool for radiotherapy centers that evaluates the impact of rescheduling decisions and their sequence on overall system performance. We consider the three main rescheduling decisions that can be made to handle machine breakdown in a radiotherapy center.

- **Delay patients (d):** Postpone the first treatment session for patients that haven't started their treatment yet.
- **Overbook patients (w):** Book two patients during the same time slot.
- **Use overtime (o):** Schedule patients using the capacity buffer provided by overtime.

Deciding to overbook has a direct impact on the resulting overtime and patient waiting time. To measure patient waiting time, we calculate the treatment lateness of patient who is scheduled to receive treatment on the previous slot. Overbooking has a larger impact on waiting times than overtime, especially when scheduled slot lengths are longer than the duration of patient treatment (because the delay has been compensated for).

Figure 6.1 summarizes all possible priority sequences for the three rescheduling decisions. For example, the application of the first priority sequence P_{odw} means that the administrators in the radiotherapy center gives the highest importance to reducing technologist overtime. Minimizing the number of the delayed patients is second in the hierarchical priority sequence, and patient overbooking is considered to have the lowest negative impact on solution quality.

6.4.2 Heuristic

We present six simple heuristics that follow the rescheduling framework (Figure 6.1). In each heuristic, we follow the inverse direction of the rescheduling decisions priority sequence. For example, to apply the heuristic with priority sequence P_{wod} (Figure 6.2), we delay patients, then we use overtime, and then we use overbooking. To reschedule by delaying patients, we use Algorithm 1. We start by listing the canceled patients due to machine breakdown. We go through the list. For each patient, we determine the next day of his last treatment session, then, we select all scheduled patients during this day that haven't started treatment yet, and they have an empty slot at the end of their treatment. We postpone the patient with the latest due dates for booking the canceled session. When we reach the determined upper bound on the number of delayed patients, we start to schedule patients using overtime

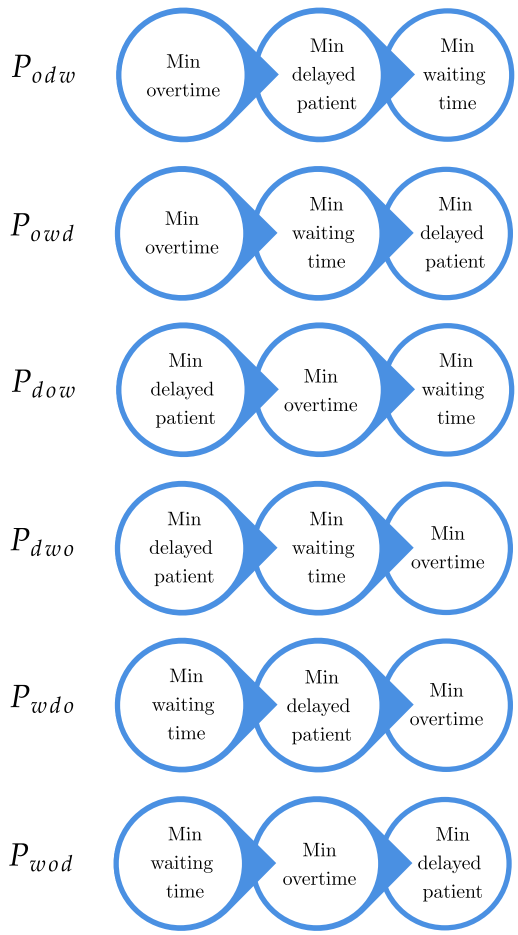

Figure 6.1 Rescheduling framework

until the maximum capacity is reached (Algorithm 2). Finally, we use overbooking until we reach the maximum total patient waiting time and overtime (Algorithm 3).

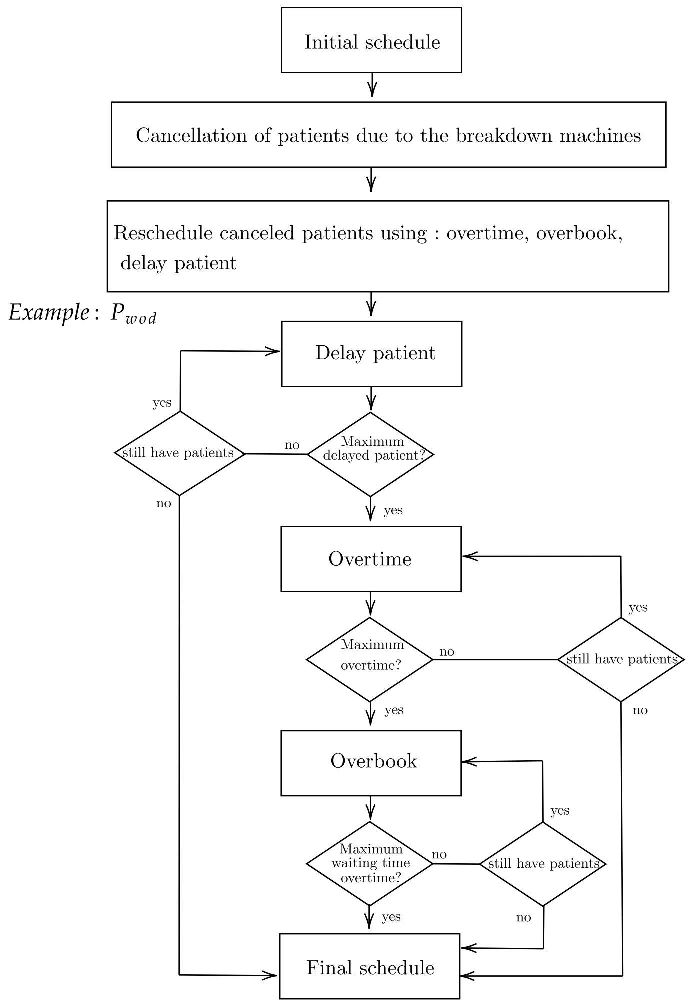

Figure 6.2 Example heuristic

Algorithm 1: Rescheduling by delaying patients

Input: Initial schedule

Remove canceled patients already scheduled by other rescheduling algorithm
let P be the maximum allowed number of delayed patients
for *i in remaining cancelled patients* **do**
    if *the total number of delayed patients* $< P$ **then**
        let day_i be the next day to the last treatment of patient i
        list patients scheduled during day_i in $list_d$
        remove patients already started their treatment from $list_d$
        for p *in* $list_d$ **do**
            let day_p be the next day to the last treatment of patient p
            S_{empty} is the set of empty slots during day_p
            if $S_{empty} = 0$ **then**
                remove p from $list_d$
            end
        end
        sort $list_d$ in descending order of patient due date
        select i' the first patient in $list_d$
        s_i' is the allocated slot to patient i' during day_i
        m_i' is the allocated machine to patient i' during day_i
        schedule i in s_i' and m_i' during day_i
        schedule i' in the first empty slot during day_p
    end
end

Output: Updated schedule

Algorithm 2: Rescheduling by using overtime

Input: Initial schedule

```
Remove canceled patients already scheduled by other rescheduling algorithm
for i in remaining canceled patients do
    let j_1 be the day of the machine breakdown
    for m in machines do
        for s in overtime slots do
            if no patient scheduled in s and m during j_1 then
                scheduled i in s and m during j_1
                calculate the overtime
                if the overtime exceed the maximum value then
                    remove i from s and m during j_1
                else
                    Move on to the next patient
                end
            end
        end
    end
end
```

Output: Updated schedule

Algorithm 3: Rescheduling by overbooking

```
Input: Initial schedule
Remove canceled patients already scheduled by other rescheduling algorithm
for i in remaining canceled patients do
    let j_1 be the day of the machine breakdown
    list patients scheduled during j_1 in list_o
    sort list_o in the ascending order of the patient lateness
    for p in list_o do
        let s_p be the slot of patient p during j_1
        let m_p be the machine of patient p during j_1
        if there is no overbooking in s_p and m_p during j_1 then
            schedule i in s_p and m_p during j_1
            calculate the waiting time and the overtime
            if the waiting time and the overtime exceed the maximum value then
                remove the patient i from s_p and m_p during j_1
            else
                Move on to the next patient
            end
        end
    end
end
Output: Updated schedule
```

Wang *et al.* (2017), Tasgetiren *et al.* (2007), and Wang et Tang (2012) all confirm that using local search for scheduling problems leads to good results. In order to enhance the quality of our solutions, we use three of the most popular local search operators: swap, insert, and inverse. After rescheduling the patients based on the initial heuristic defined above, we apply the local search algorithm on the list of canceled patients, excluding delayed patients (Algorithm 4). In fact, we randomly select two canceled patients, and the local search operator to apply. We keep the new solution if performance is improved. Since we only consider the list of canceled patients moved during the breakdown day, the only impacted performance indicators are overtime and waiting time. For each two candidate patients, we carry out 1000 algorithm iterations.

Figure 6.3 is an example of our local search algorithm applying the operators to patients p_1 and p_{12} from a list of 13 patients. We define the three local search operators used in our improvement algorithm.

The swap operator requires two different patients exchange their positions in the schedule on the day the machine breaks down.

The insert operator removes a patient from their actual position to insert them directly after another chosen patient so each patient takes the position of the previous one.

The inverse operator inverses the sequence of patients placed between two positions of two chosen patients.

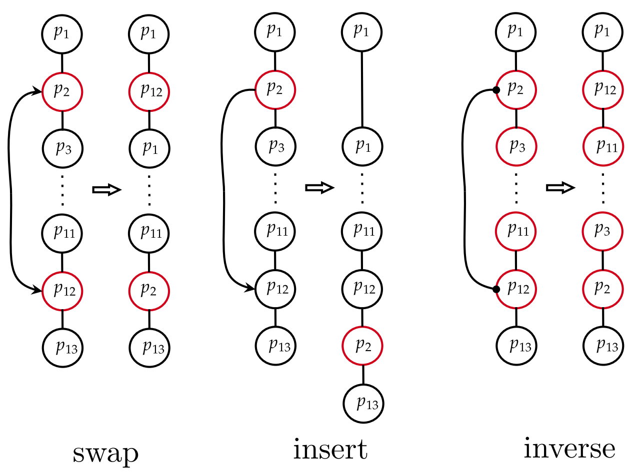

Figure 6.3 the local search operators

Algorithm 4: Local search algorithm

Input: Updated schedule after rescheduling canceled patients

Remove delayed patient from the list of canceled patients ($list_c$)
Let overtime and waiting time be the indicators to consider
Let K_1 be the first indicator to prioritize
Let K_2 be the second indicator to prioritize
Let V_1 be the value of K_1 after rescheduling
Let V_2 be the value of K_2 after rescheduling
$Iterations_1 = 0$
$Iterations_2 = 0$
while $Iterations_1 \leq 1000$ **do**
    $Iterations_1 = Iterations_1 + 1$
    Select randomly 2 patients (P_1, P_2) from $list_c$
    Select randomly one local search operator : Swap(P_1, P_2), Insert(P_1, P_2), Inverse(P_1, P_2)
    Apply the selected local search operator
    if V_1 *is improved* **then**
        Update the value of V_1
        Update the schedule after applying the selected local search operator
    end
end
while $Iterations_2 \leq 1000$ **do**
    $Iterations_2 = Iterations_2 + 1$
    Select randomly 2 patients (P_1, P_2) from $list_c$
    Select randomly one local search operator : Swap(P_1, P_2), Insert(P_1, P_2), Inverse(P_1, P_2)
    Apply the selected local search operator
    if V_1 *is not deteriorated and* V_2 *is improved* **then**
        Update the value of V_1 and V_2
        Update the schedule after applying the selected local search operator
    end
end

Output: Updated schedule after the application of the local search

6.4.3 Model

The initial schedule represents the preliminary radiotherapy assignment based on the predefined center appointment grid. The binary parameters X_{is}^{km} are equal to 1 if a patient $i \in I$ was assigned to machine $m \in M$, slot $s \in S$, and day $k \in K$, depending on their treatment time t_i. The patient $i \in I$ is assigned to the slot s of adequate length (d_s^{km}), on the machine m, and the day k. Each patient $i \in I$ has to perform a determined number of treatment sessions e_i before their due date f_i.

When the radiotherapy machine $m \in M$ is broken down during the slot $s \in S_D$, and day $k \in K_D$, the binary parameters b_{skm} take the value 1. The rescheduling of canceled patients $i \in I_c$, including those in the waiting room $i \in I_w$, require one of the three rescheduling decisions: delay patients, overbook patients, use overtime. To delay patients, we have to exclude the set of patients that can't change their treatment schedule $i \in I_f$, and to respect the maximum number of delayed patients p. While overbooking patients, we have to consider the maximum patient waiting time allowed on machine m, day k (w_{km}). The set of slots is divided into two sets: the set of slots not reserved for overtime S_{no}, and the set of slots reserved for overtime S_o. There are also days $k \in K_{nw}$ and slots $s \in S_{nw}$ where overbooking is forbidden. So, in order to use overtime, it must be allowed given the day and slots, and the ability of machine m to work overtime on day k (a_{km}). Moreover, we can't exceed the maximum overtime allowed on machine m, and day k (o_{km}).

Our model decides:

- if patient i is assigned to machine m, slot s and day j, using binary variables x_{is}^{km}
- if patient i starts their treatment on machine m, slot s, day j, using binary variables y_{is}^{km}
- if patient i is moved from slot s, machine m, day k, using binary variables α_{is}^{km}
- if patient i is delayed to day k, using binary variables β_i^k

The model measures:

- patient treatment lateness on slot s, machine m, day k, using variables δ_s^{km}
- overtime due to treatment lateness on machine m, day k, using variables γ^{km}
- total overtime on machine m, day k, using variables ϕ^{km}

The proposed model below accounts for all feasibility constraints related to patient scheduling. Constraints (1)-(3) represent the unicity constraints. Constraints (1) require each patient to start their treatment only once. Constraints (2) limit the number of slots attributed to a patient per day to one. Constraints (3) ensure that if overbooking is not allowed, at most one patient can be assigned to a slot, machine, and day. In contrast, constraints (4) permit the assignment of two patients to the same slot, day, and machine. Constraints (5) ensures that all patients start treatment before their due date.

$$\sum_{k\in K}\sum_{s\in S}\sum_{m\in M} y_{is}^{km} \leq 1, \quad \forall\, i \in I \tag{1}$$

$$\sum_{s\in S}\sum_{m\in M} x_{is}^{km} \leq 1, \quad \forall\, i \in I,\, k \in K \tag{2}$$

$$\sum_{i\in I} x_{is}^{km} \leq 1, \quad \forall s \in S_{nw},\, k \in K_{nw},\, m \in M \tag{3}$$

$$\sum_{i\in I} x_{is}^{km} \leq 2, \quad \forall\, s \in S,\, k \in K,\, m \in M \tag{4}$$

$$\sum_{k=1}^{f_i}\sum_{s\in S}\sum_{m\in M} y_{is}^{km} \geq 1, \quad \forall\, i \in I \tag{5}$$

Constraints (6)-(8) apply the patient treatment plan specifications. A patient cannot receive more than the prescribed radiotherapy sessions (Constraints (6)), which have to start once on a given machine and slot (Constraints (7)), and have to be performed daily in a consecutive way (Constraints (8)).

$$\sum_{k\in K}\sum_{s\in S}\sum_{m\in M} x_{is}^{km} \leq e_i, \quad \forall\, i \in I \tag{6}$$

$$\sum_{s\in S}\sum_{m\in M} x_{is}^{hm} \leq \sum_{s\in S}\sum_{m\in M} M(1-y_{is}^{km}), \quad \forall\, i \in I,\, k \in K,\, h \in K,\, h < k \tag{7}$$

$$\sum_{s\in S}\sum_{m\in M} y_{is}^{km} \leq \sum_{s\in S}\sum_{m\in M} x_{is}^{km}, \quad \forall\, i \in I, k \in K,\, h \in \{k, .., min(k+e_i-1, |K|)\} \tag{8}$$

When a machine breaks down, Constraints (9) permit canceling patients from the slots and days where the breakdown occurs. Constraints (10) and Constraints (11) prevent rescheduling patients that can't change their treatment schedule. Constraints (12) ensure same day treatment for patients that are in the waiting room during the period of machine breakdown.

$$b_{skm} x_{is}^{km} = 0, \quad \forall\, i \in I,\, s \in S,\, k \in K,\, m \in M \tag{9}$$

$$X_{is}^{km} \leq x_{is}^{km}, \quad \forall i \in I_f,\, s \in S,\, k \geq 2,\, m \in M \tag{10}$$

$$X_{is}^{1m} \leq x_{is}^{1m}, \quad \forall i \in I_f,\, s \in S,\, m \geq 2, \tag{11}$$

$$\sum_{s\in S}\sum_{k\in K_D}\sum_{m\in M} x_{is}^{km} \geq 1 \quad \forall\, i \in I_W \tag{12}$$

Constraints (13)-(17) concern delaying patients. While rescheduling due to machine breakdown, we have two possibilities: move patients to another slot or machine on the same day, or delay their treatment to another day. This condition is guaranteed by constraints (13). Constraints (14) determine if the patient is delayed, and if not, constraints (15) confirm that they are moved and will be treated the same day. Constraints (16) and constraints (17) require the upper bound on the number of moved patients and the number of delayed patients be respected.

$$\alpha_{is}^{km} + \beta_i^k \leq 1, \quad \forall i \in I,\, s \in S\, k \in K,\, m \in M \tag{13}$$

$$\sum_{s \in S} \sum_{m \in M} x_{is}^{km} \geq \sum_{s \in S} \sum_{m \in M} X_{is}^{km} - \beta_i^k, \quad \forall\, i \in I,\, k \in K \tag{14}$$

$$x_{is}^{km} \geq X_{is}^{km} - \alpha_{is}^{km} - \beta_i^k, \quad \forall i \in I,\, s \in S\, k \in K,\, m \in M \tag{15}$$

$$\sum_{i \in I} \sum_{k \in K} \sum_{s \in S} \sum_{m \in M} \alpha_{is}^{km} \leq |I_c|, \tag{16}$$

$$\sum_{i \in I} \sum_{k \in K} \beta_i^k \leq P, \tag{17}$$

Constraints (18)-(21) regard patient waiting time. Constraints (18)-(20) calculate the patient treatment lateness for a given slot, which represents the waiting time of the patient scheduled for the next slot of the same machine. Constraints (21) enforce the maximum waiting time per machine and day.

$$\delta_s^{km} \leq d_s^{km} - \sum_{i \in I} \sum_{s \in S_{ov}} t_i x_{is}^{km} + \delta_{s-1}^{km}, \quad \forall s \in S_{no}, s \geq 2,\, k \in K,\, m \in M \tag{18}$$

$$\delta_1^{km} \leq d_1^{km} - \sum_{i \in I} \sum_{s \in S_{ov}} t_i x_{i1}^{km}, \quad \forall\, k \in K,\, m \in M \tag{19}$$

$$\delta_s^{km} \leq 0, \quad \forall\, s \in S,\, k \in K,\, m \in M \tag{20}$$

$$\sum_{s=1}^{|S|-1} \delta_s^{km} \leq w^{km}, \quad \forall\, k \in K,\, m \in M \tag{21}$$

Constraints (22)-(26) pertain to overtime constraints. Constraints (22) permit assigning a patient to overtime slots if it is allowed in a given day and machine. Constraints (23)-(25) determine the overtime value per day and machine; it is the sum of the lateness off the last patient with the treatment time of the patients allocated to the overtime slots. Constraints (26) limit the maximum overtime per day and machine.

$$x_{is}^{km}(1 - a_{km}) = 0, \quad \forall\, i \in I,\, s \in S_o,\, k \in K,\, m \in M \tag{22}$$

$$\gamma^{km} \geq -\delta_{|S_{no}|}^{km}, \quad \forall\, k \in K,\, m \in M \tag{23}$$

$$\gamma^{km} \geq 0, \quad \forall\, k \in K,\, m \in M \tag{24}$$

$$\phi^{km} = \sum_{i \in I} \sum_{s \in S_o} t_i x_{is}^{km} + \gamma^{km}, \quad \forall k \in K,\, m \in M \tag{25}$$

$$\phi^{km} \leq o^{km}, \quad \forall\, k \in K,\, m \in M \tag{26}$$

Constraints (27)-(30) define the domain of the variables.

$$x_{is}^{km} \in \{0,1\}, y_{is}^{km} \in \{0,1\}, \alpha_{is}^{km} \in \{0,1\}, \quad \forall\, i \in I,\, s \in S,\, k \in K,\, m \in M \tag{27}$$

$$\beta_i^k \in \{0,1\}, \quad \forall\, i \in I,\, k \in K \tag{28}$$

$$\delta_s^{km} \in \mathbb{R}, \quad \forall\, s \in S,\, k \in K,\, m \in M \tag{29}$$

$$\gamma^{km} \in \mathbb{R}, \phi^{km} \in \mathbb{R}, \quad \forall\, k \in K,\, m \in M \tag{30}$$

Our objective function aims to minimize three elements: the number of delayed patients (equation (31)), patient waiting time (equation (32)), and machine overtime (equation (33)).

$$\text{The number of delayed patients} \quad \omega_1 = \sum_{i \in I} \sum_{k \in K} \beta_i^k \tag{31}$$

$$\text{Waiting time} \quad \omega_2 = -\sum_{s=1}^{|S|-1} \sum_{k \in K} \sum_{m \in M} \delta_s^{km} \tag{32}$$

$$\text{Overtime} \quad \omega_3 = \sum_{k \in K} \sum_{m \in M} \phi^{km} \tag{33}$$

We adopt the hierarchical optimization approach to solve the problem, so the resolution is performed by steps. We minimize objective by objective and the optimal solution obtained is included as an upper bound among the constraints. Figure 6.4 presents an example of hierarchical optimization based on priority sequence P_{wod}. We start by minimizing waiting time by considering constraints (1)-(30). The optimal solution obtained represents an upper bound on total waiting time while minimizing overtime in the second step. The third step consists of minimizing the number of delayed patients by considering the constraints (1)-(30) together with the maximum overtime and maximum waiting time obtained during the first two steps.

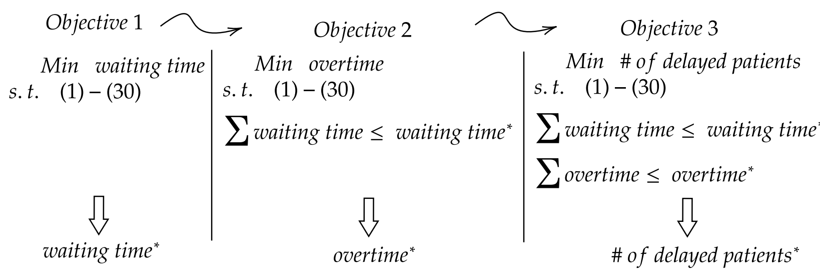

Figure 6.4 Example of hierarchical optimization

6.5 Experiments and results

We propose a case study to illustrate our approach. We collaborate with the CICL. Their radiotherapy center contains four linacs with the same characteristics, so patient treatment can be performed on any machine. The center adopts the appointment grid system for scheduling, and there are two slot time categories: 15 min and 20 min.

6.5.1 Current approach

The CICL doesn't have a predefined procedure for handling machine breakdown. On a day with machine failure, technologists may be forced to work the maximum possible overtime. Moreover, the CICL may overbook patients based on their experience estimating real treatment time. The CICL delays patients only as a last resort.

6.5.2 Experimental design

In our experimental design we present several scenarios for evaluating the proposed approach under different working situations.

Our approach is based on updating the initial schedule. Therefore, using the heuristic or the optimization model, we start from the previously existing schedule. To instantiate the initial schedule for our experimentation, we book fictive patients to predetermined CICL appointment grids, respecting the consecutive treatment sessions, and the need for assignment to the same slot and machine. Our schedule horizon is four weeks. We classify patients into their slot categories. We generate the number of treatment sessions for each patient, their due date, and their treatment time randomly from real data provided by the CICL, and listed based on patient slot category.

We allow utilization of overbooking and overtime only on the day of machine breakdown. We set the maximum additional patient waiting time relative to the initial situation at 45 min per day and machine. Moreover, the additional overtime per day and machine may not exceed 20% of the machine breakdown duration. The maximum number of delayed patients equals 40 % of the number of canceled patients. At the moment of machine breakdown, we assume we have two patients in the waiting room that have to be treated on the same day. The fill rate of the schedule over the four weeks takes three levels: 75%, 85%, 95%.

Table 6.1: Predicted and updated machine breakdown duration

Predicted duration	Updated duration
2h	2h
	4h
	6h
4h	4h
	6h
6h	6h

We assume that machine breakdown duration is predictable, and then it is updated. For the predictable and the updated, we test three duration values: 2h, 4h, 6h. Table 6.1 summarizes all the combinations that we include in our experimentation. To apply the proposed model and the heuristic under this condition, we follow the process described in figure 6.5. When the machine breaks down, we use our rescheduling approach based on the predicted breakdown duration, and we generate a new schedule by dispatching all canceled patients. If the predicted duration is not updated, and the machine is repaired, the generated schedule is the final solution. If the machine is still broken down, we update the breakdown duration, and reapply the rescheduling approach by preserving the decisions that were already made during the previous workday period.

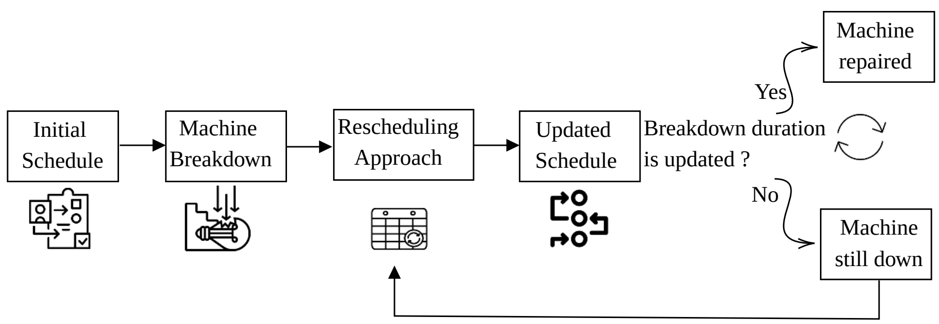

Figure 6.5 Rescheduling process

6.5.3 Performance indicators

The obtained solutions were evaluated based on three performance indicators: number of delayed patients, patient waiting time, and used overtime.

Number of delayed patients The total number of patients whose first treatment day was postponed without disrespecting their due date.

Patient waiting time Patient waiting time consists of direct and indirect waiting time. In this article we only account for direct patient waiting time, which is determined as the positive gap between the actual and planned daily treatment start times. We define this indicator as the total additional waiting time created due to rescheduling by overbooking.

Overtime The sum of the positive difference between the real and scheduled completion times. The overtime we consider in this article is the additional overtime due to machine breakdown relative to the initial schedule.

6.5.4 Results

We evaluated the proposed rescheduling framework using the heuristic and optimization model, and experimentation settings defined in subsection 6.5.2. We implemented our model with JuMP (Dunning *et al.*, 2017) for mathematical optimization in Julia. We solved it using CPLEX Solver, and a PC with an Intel Core i7 2.80 GHZ and 16 GB RAM processor. We ran 108 scenarios by adjusting the schedule fill rate, and the predicted and updated machine breakdown duration. Ten initial schedules were tested with each of the generated scenarios by applying the heuristic and the optimization model, and by considering the six possible priority sequences for the rescheduling decisions. The computational time increases when increasing the machine breakdown duration and the schedule fill rate. The average computational time for the heuristic is about 40 seconds, and varies for the optimization model between 5.18 minutes and 11.47 minutes.

To measure the contribution of the local search algorithm on the efficiency of the rescheduling heuristic, we start by rescheduling patients following only algorithms 1, 2, and 3. Then we redo all the tests with the local search algorithm (Algorithm 4). Figure 6.6 proves that the application of the local search algorithm minimizes overtime by 25%, while wait times deteriorated by 17% for rescheduling decision sequences that prioritize overtime. However, for rescheduling decision sequences that prioritize wait times, figure 6.7 shows that applying the local search algorithm enhances wait times by 17%, while overtime is almost the same. We only consider the rescheduling heuristic with the local search algorithm in all of the following analysis.

Figures 6.12- 6.20 (Appendix) summarize the results of the proposed model and heuristic following different rescheduling decision priorities and adjusting the schedule fill rate and breakdown duration. Afterwards, we present a more detailed analysis to evaluate the impact of each factor.

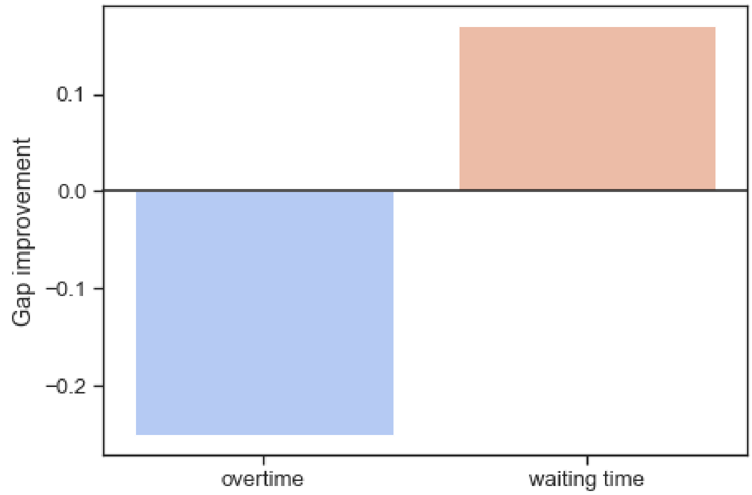

Figure 6.6 The gap improvement using the local search algorithm and prioritizing overtime

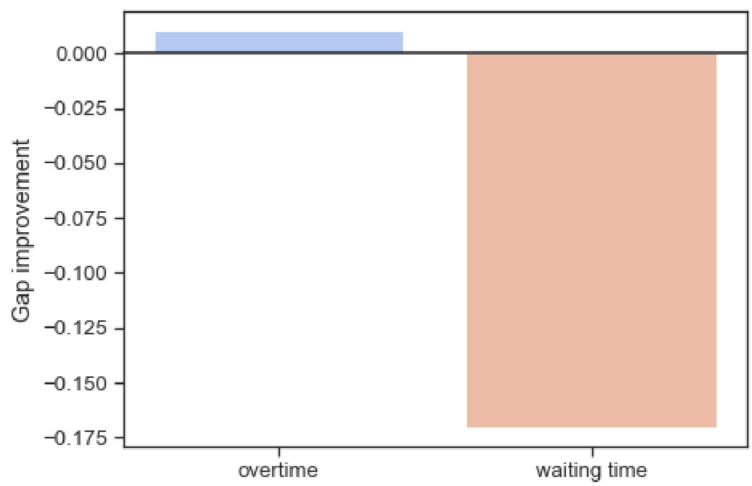

Figure 6.7 The gap improvement using the local search algorithm and prioritizing wait times

Figures 6.12- 6.20 (Appendix) shows us that in general the mathematical models outperform the heuristics. To get more accurate information, we calculate the difference between values in the two methods for each performance indicator by considering their priority levels. Figure 6.8 confirms that the heuristic works very well in terms of the number of delayed patients for any priority sequence of the rescheduling decisions. Moreover, figure 6.8 shows that the heuristic gives good results for the first indicator prioritizes. However, the difference is great when overtime or waiting time is the second indicator prioritized; the median of the difference is about 50 min and 25 min for the overtime and waiting time respectively.

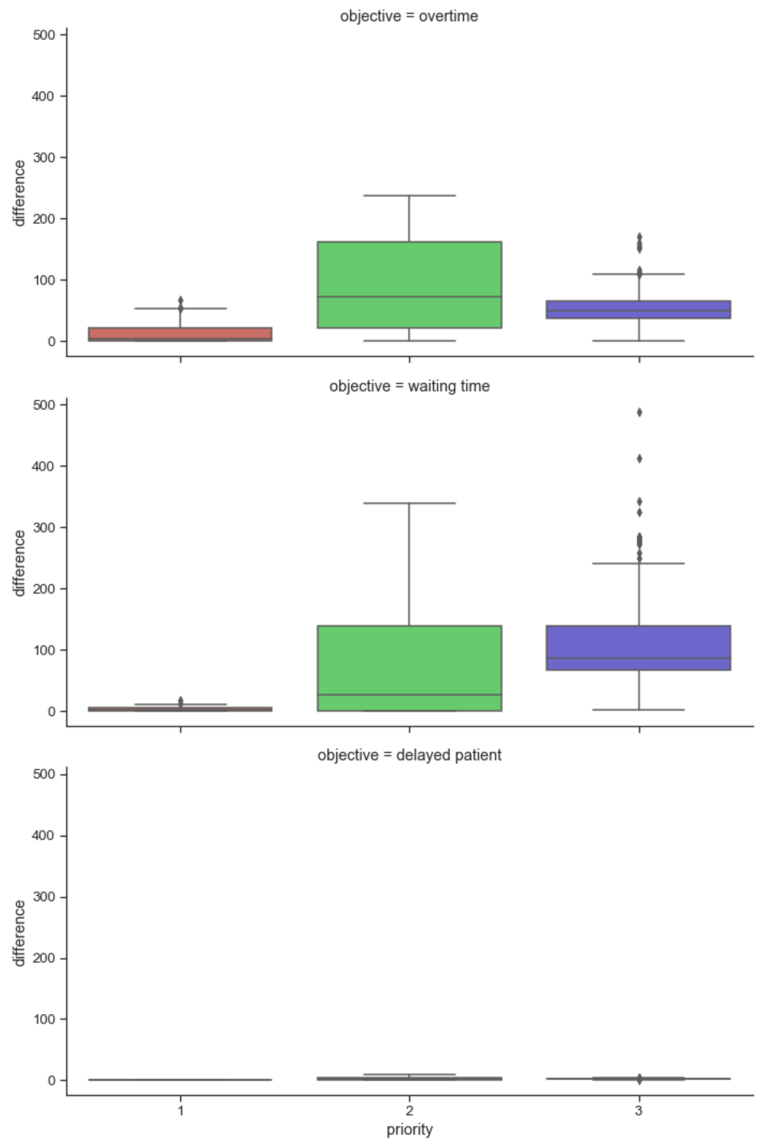

Figure 6.8 The difference in performance indicator values using both the heuristic and the model

We perform a statistical analysis on the obtained results to study the interaction between the indicators, and to evaluate the impact of the experimentation factors. The results of the Spearman's correlation test (Figure 6.9) confirm that overtime and waiting time are strongly correlated. We also apply the Friedman non-parametric statistical test on our data, taking into account our performance indicators, and considering the data variation due to our experimentation factors: the priority sequence of the rescheduling decisions, the breakdown duration, and the schedule fill rate. The Friedman test shows us that all factors are significant, with a p-value equal to 0.

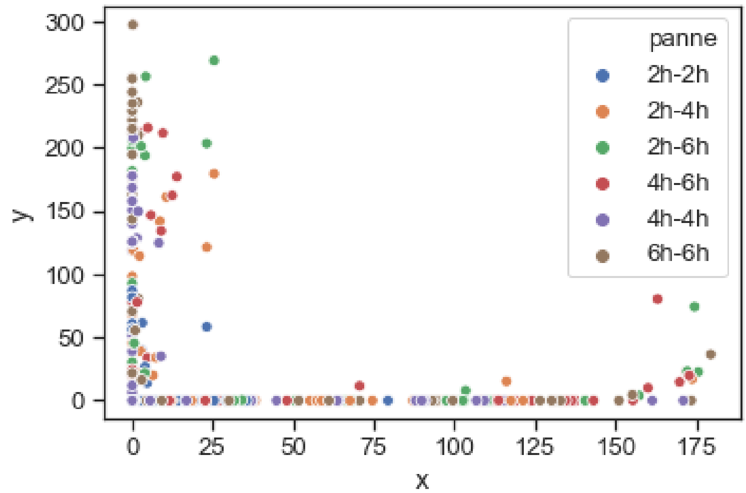

Figure 6.9 the Spearman's correlation between overtime and wait times

Tables 6.2- 6.4 present the average of each performance indicator by varying the priority sequence, the breakdown duration, and the schedule fill rate. The obtained values confirm the Friedman test results; we have a significant difference between performance when varying the experimentation factors.

Machine breakdown duration directly influences the number of delayed patients, patient waiting times, and overtime. Long breakdown durations deteriorate the value of the performance indicators. Moreover, Table 6.2 shows that accurate prediction of breakdown duration slightly improves solution quality. Furthermore, we conclude from table 6.3 that the more patients we schedule initially, the more we increase wait times and overtime while rescheduling.

Table 6.2: The average of the performance indicators by predicted and updated machine breakdown duration

Predicted - Updated	Overtime	Waiting time	Number of delayed patients
2h-2h	18.42	13.43	0.75
2h-4h	52.42	35.84	1.75
2h-6h	83.03	50.01	2.98
4h-4h	52.81	33.04	1.66
4h-6h	81.15	53.71	3.07
6h-6h	79.64	46.30	2.75

Table 6.3: The average of the performance indicators by schedule fill rate

Fill rate	Overtime	Waiting time	Number of delayed patients
95%	68.28	50.71	2.32
85%	61.08	36.49	2.25
75%	54.37	28.96	1.92

Performance of each indicator follows the decision prioritization sequence. The best results for each performance indicator are obtained when the corresponding rescheduling decision is first in the prioritization sequence. For example, the sequence that prioritizes minimization of the number of delayed patients (P_{dow}, P_{dwo}), avoids delaying even one patient in all scenarios. We also find that when prioritizing the number of delayed patients (P_{dow}, P_{dwo}) yields good results in the second indicator considered. Based on table 6.4, the average overtime of the sequences (P_{odw}, P_{owd}, P_{dow}) is approximately equal to the average waiting time of the sequences (P_{dwo}, P_{wdo}, P_{wod}), about 2 min. However, the average waiting time of the sequences that prioritize overtime with respect to wait times (P_{odw}, P_{owd}, P_{dow}) is 75.45 min, and the average overtime of sequences that prioritize wait times with respect to overtime (P_{dwo}, P_{wdo}, P_{wod}) is 120 min. We have satisfactory results for overtime and wait times by reporting six patients in the average with the sequences (P_{owd}, P_{wod}).

Table 6.4: The average of the performance indicators by the rescheduling priority sequence

Priority sequence	**Overtime**	**Waiting time**	**Number of delayed patients**
P_{odw}	0.25	92.64	0.94
P_{owd}	0.36	28.06	5.74
P_{dow}	6.12	105.65	0.0
P_{dwo}	180.80	2.74	0.0
P_{wdo}	152.74	2.28	0.38
P_{wod}	27.18	0.95	5.92

6.6 Discussion and managerial insights

In this section, we highlight managerial insights to guide healthcare administrators in the use of our rescheduling decision tool and the provided analysis of our results.

We suggested different rescheduling strategies based on the priority sequence of the main rescheduling decisions, and we presented the results in each case to give a view of how that decision affects performance. In the present analysis, we want to further clarify the interaction between the indicators to help choose the best overall rescheduling strategy. For a better data visualization, we use Min-Max scaling for data transformation in order to scale data to a fixed range between 0 and 1. We transform the obtained values for overtime, wait times, and the number of delayed patients using equation 6.1.

$$z = \frac{x - min(x)}{max(x) - min(x)} \tag{6.1}$$

where z : the normalized value from $min - max$ scaling;

x : the original value;

Figure 6.10 compares our three objectives using a unified scale. This allows us to choose the best rescheduling strategy by accounting for all of the performance indicators at the same time. Managers could consider the presented figure to be historical data related to

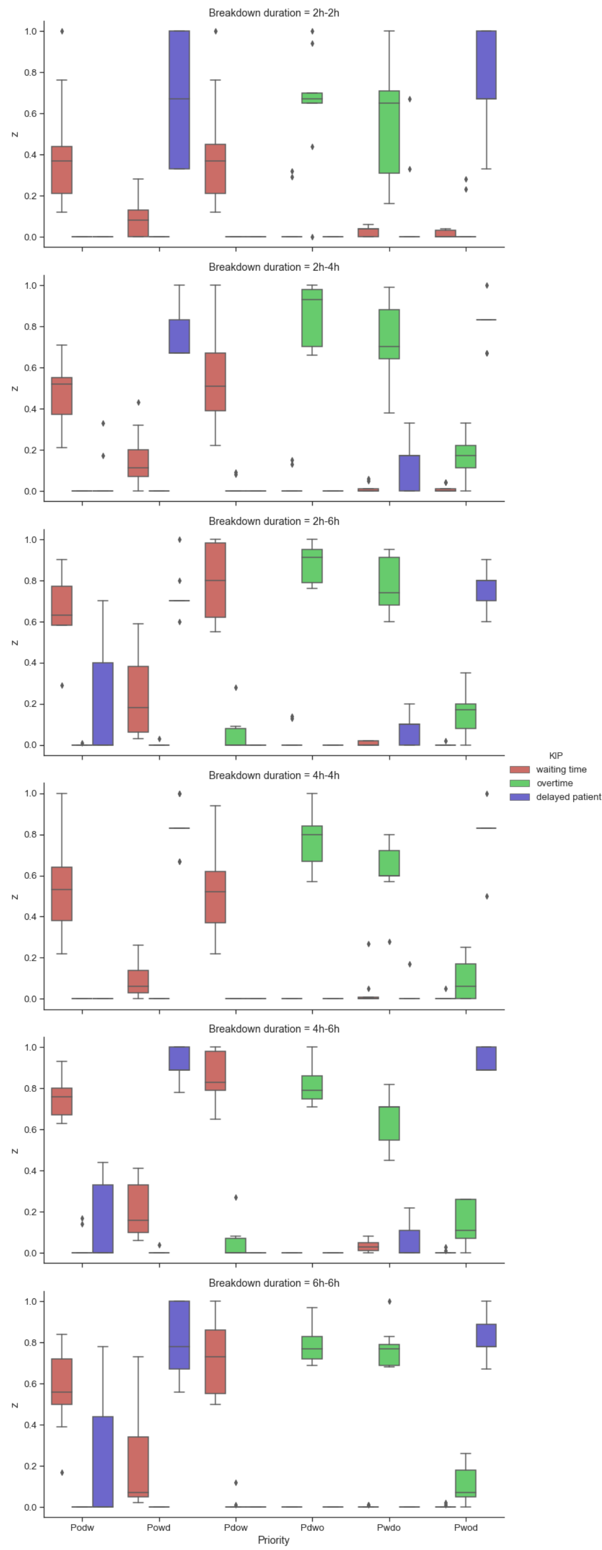

Figure 6.10 The $min - max$ scaling data transformation

machine breakdown, and select a suitable solution according to their internal objectives. For example, when predicted machine breakdown duration is 2h, and the actual duration is 4h (Figure 6.11), if the managers prioritize overtime, the best priority sequences are P_{odw}, P_{owd}, and P_{dow}. If they prefer to minimize wait times with respect to the number of delayed patients, P_{owd} is the best strategy to adopt. In the opposite case, or if they give the same importance to wait times as the number of delayed patients, the best priority sequence is P_{odw}. When prioritizing wait times, if we classify overtime and the number of delayed patients at the same priority level, or we promote minimization of the number of delayed patients, we find that P_{dwo} is the best priority sequence. However, if we give equal priority to all performance indicators, P_{odw} is the best solution to apply because the score obtained by the $min - max$ scaling is equal to 0.5 for the waiting time and zero for the other indicators.

This representation will help managers by giving them an overview of overall system performance, and guide them to choose the best rescheduling strategy, or adjust their priorities if they penalize the other indicators in an exaggerated way.

We could also take advantage of the present study by proposing the best predicted machine breakdown duration to reschedule patients based on the obtained results and historical data. In this article we assume the predicted breakdown duration could be updated with the actual breakdown duration. The results vary depending on this information. In our case study we don't have data related to breakdown duration, nor the reliability of expected repair duration in real life. To demonstrate our second managerial insight, we present an example where the predicted breakdown duration equals four hours. We assume that breakdown

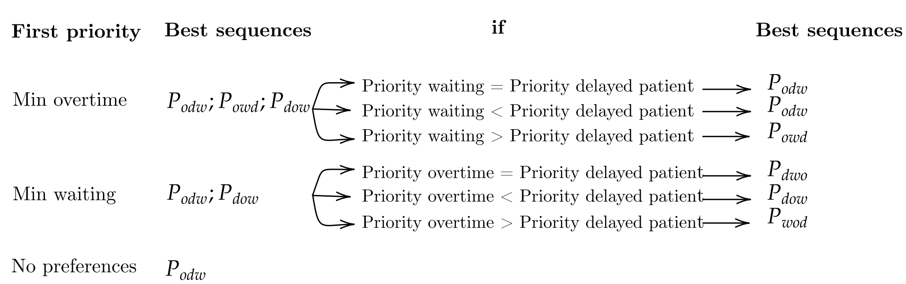

Figure 6.11 Example of optimal rescheduling strategy selection

duration couldn't exceed six hours. We consider two possibilities here. The first possibility is to trust the predicted duration and reschedule only patients covered the four hours of machine breakdown, and if that duration is updated to six hours, reschedule the booked patients during the remaining two hours. The second possibility is to assume the worst case scenario from the beginning, and to reschedule patients by taking into account the maximum breakdown duration. To compare the two possibilities, we would have to consider the probability of each case occurring. Table 6.5 gives the average waiting time with respect to a fill rate of 85%, and the strategies that prioritize the overtime indicator. We calculate the values related to the predicted breakdown duration of four hours and six hours. Table 6.6 gives us the expected wait times via random probabilities presented in the table and applying equation (6.2). When the probability to repair the machine in four hours is lower tan 75%, the expected wait times become longer than the average wait times if we assume that there will be six hours of the machine breakdown in the beginning. So, we recommend assuming at the beginning that all patients within the possible range of machine breakdown be rescheduled.

$$Expected\ waiting\ time = Pr(4h-4h) * Awt(4h-4h) + Pr(4h-6h) * Awt(4h-6h) \quad (6.2)$$

Table 6.5: Average wait times by machine breakdown duration

	Predicted-Actual		
	4h-4h	**4h-6h**	**6h-6h**
Average wait times (Awt)	67.30	104.12	76.03

Table 6.6: Expected wait times by breakdown duration probability

Pr(4h-4h)	**Pr(4h-6h)**	**Expected wait times**
95	5	68.85
90	10	70.70
85	15	72.55
80	20	74.40
75	25	76.25
70	30	78.10

We deduce that the accurate breakdown duration predication, and the availability of historical data, reinforce the utility of this paper, and will help lead managers to efficient decisions.

6.7 Conclusion

In this paper, we presented a rescheduling procedure for radiotherapy patients. We defined three main rescheduling decisions: delay patients, overbook patients, and use overtime. We developed an optimization model and a simple heuristic for rescheduling. We applied the proposed rescheduling approach in a well-defined rescheduling framework that includes all possible decision priority sequences. We collaborated with the CICL; we used their data concerning appointment grids and machines, and based on their historical data we generate treatment times, the number of sessions per patient, and treatment plan due dates. Like most radiotherapy centers, CICL doesn't have a well-defined and structured rescheduling procedure, and machine breakdown data is not documented. To evaluate our approach, we generated several scenarios by changing the values of predicted and updated machine breakdown duration, as well as the schedule fill rate. Comparison of the solutions is based on the number of delayed patients, wait times, and overtime. The results confirm that the efficiency of the proposed heuristic depends on the priority sequence of the rescheduling decisions. In general, we find that the optimization model outperforms the heuristic. However, the heuristic gives satisfactory results. The aim of this study is not just to evaluate our methodology, but also to present a structured rescheduling tool that will help managers at radiotherapy centers understand the impact of their decisions, and the consequence of their priority sequence on the performance of the overall system. We have provided tools to help each manager choose the solution that best meets their strategic policy.

A limitation of this study is the lack of historical data related to machine breakdown. Such information would give a clearer idea of the statistical distribution of machine breakdowns, and allow us to propose solutions more adapted to the studied case. In addition, we were unable to compare our approach with the current rescheduling procedure used in the cancer center, because they don't have a clear and well documented rescheduling process.

We aim to improve our rescheduling heuristic in the future to present a simple tool that help managers more easily achieve improvement in the performance of their rescheduling decisions.

6.8 Appendix

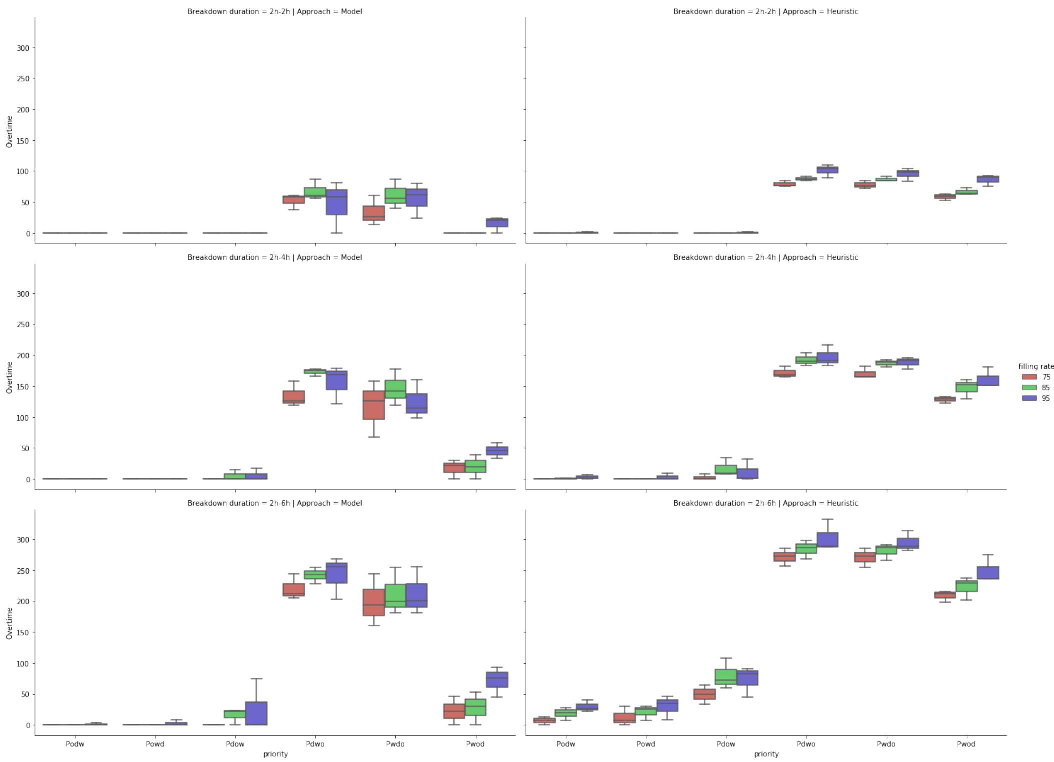

Figure 6.12 Overtime for 2h machine breakdown

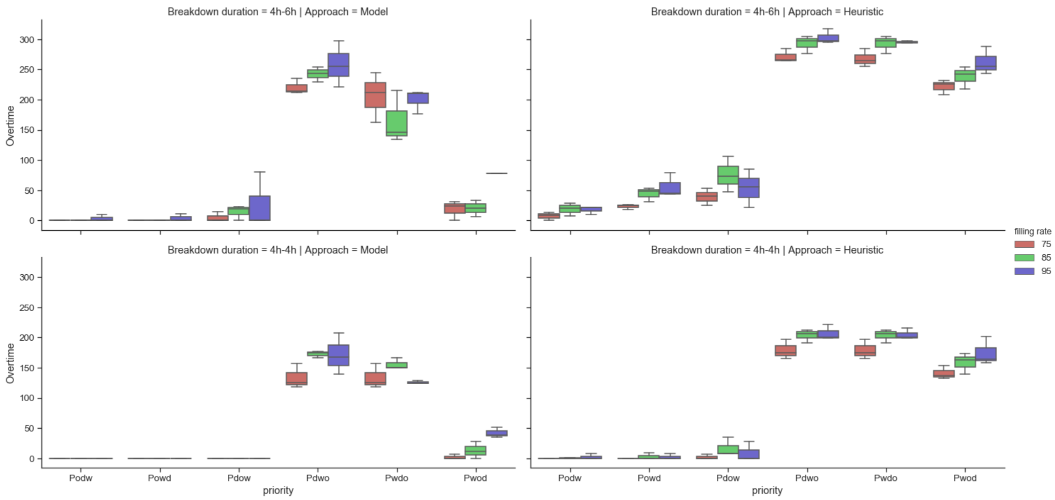

Figure 6.13 Overtime for 4h machine breakdown

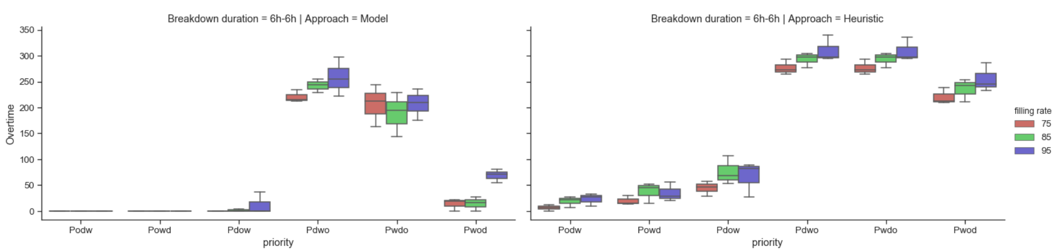

Figure 6.14 Overtime for 6h machine breakdown

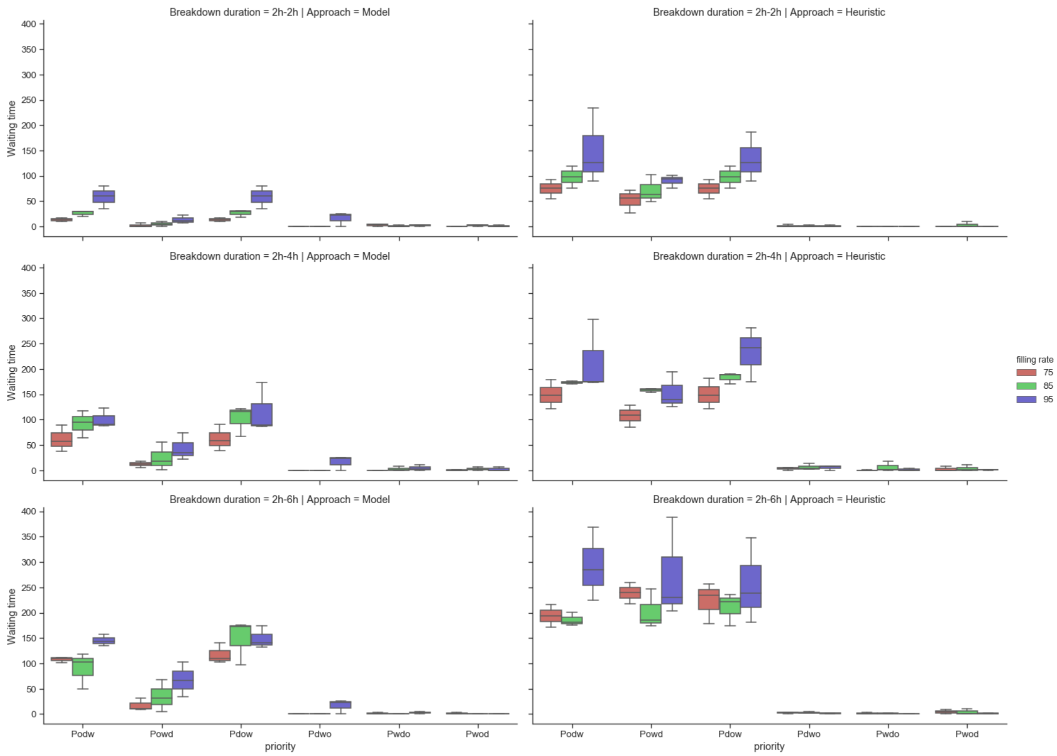

Figure 6.15 Wait times for 2h machine breakdown

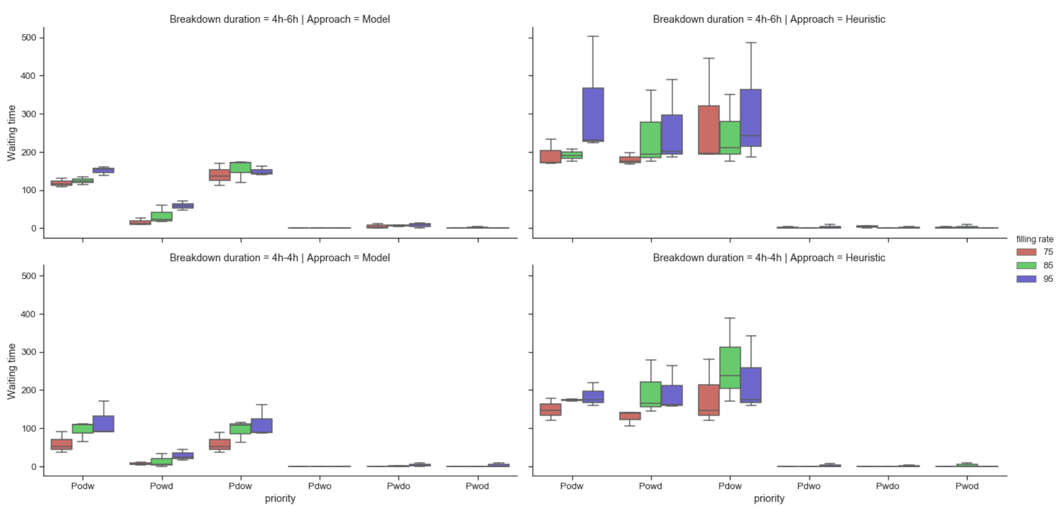

Figure 6.16 Wait times for 4h machine breakdown

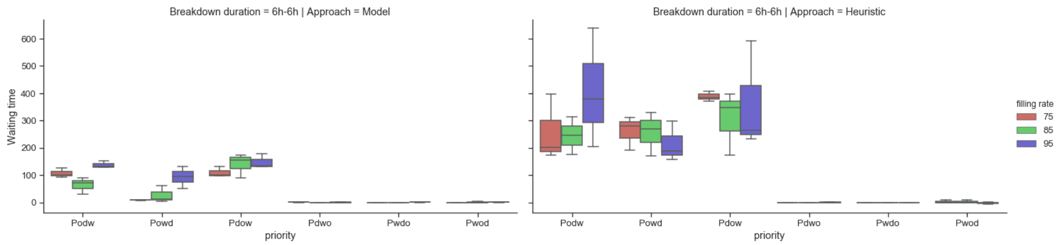

Figure 6.17 Wait times for 6h machine breakdown

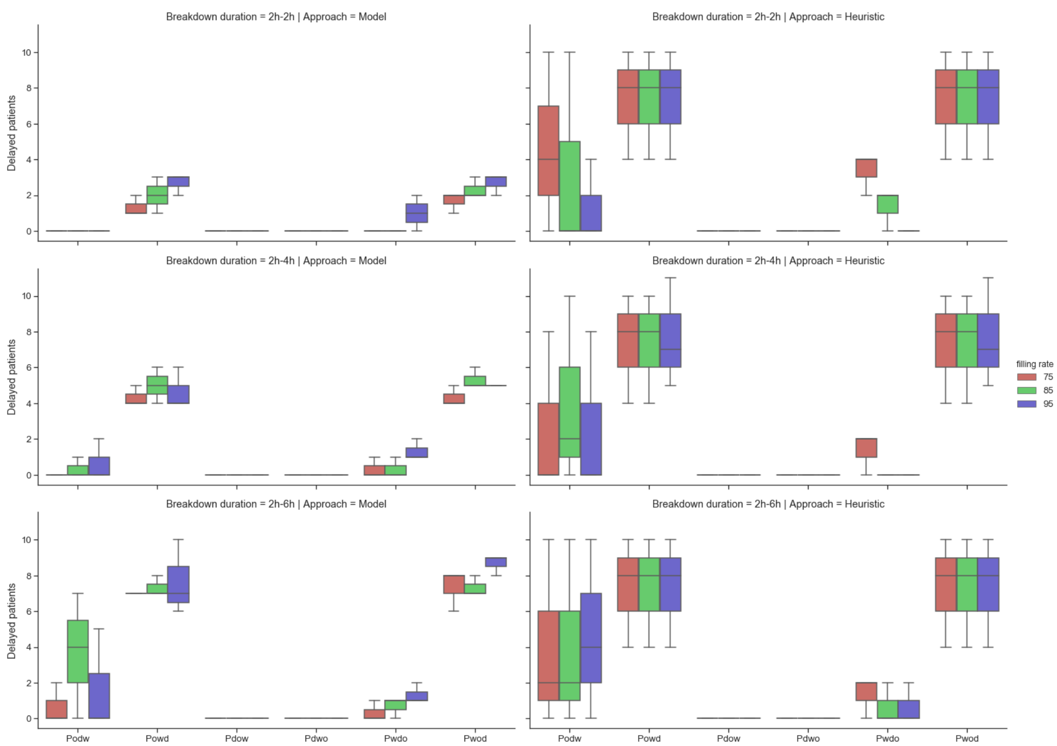

Figure 6.18 Number of delayed patients for 2h machine breakdown

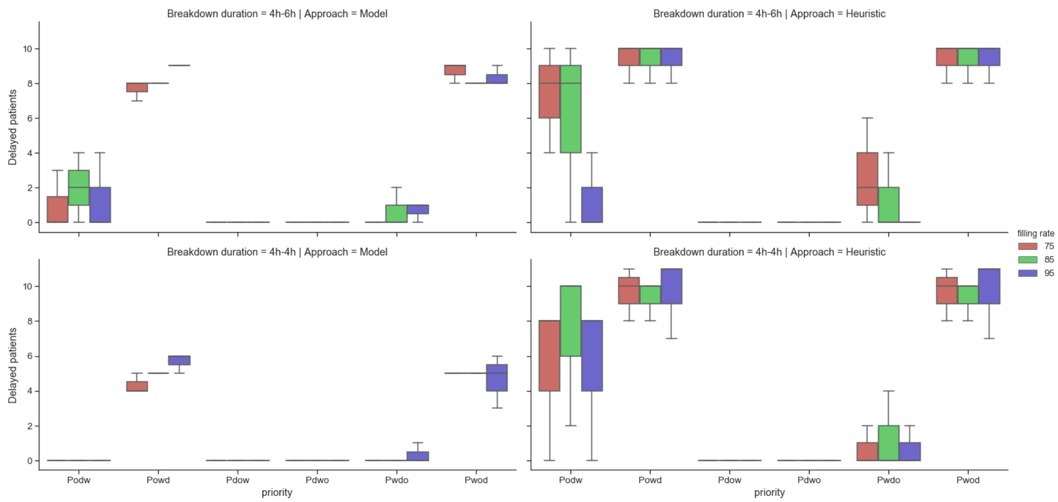

Figure 6.19 Number of delayed patients for 4h machine breakdown

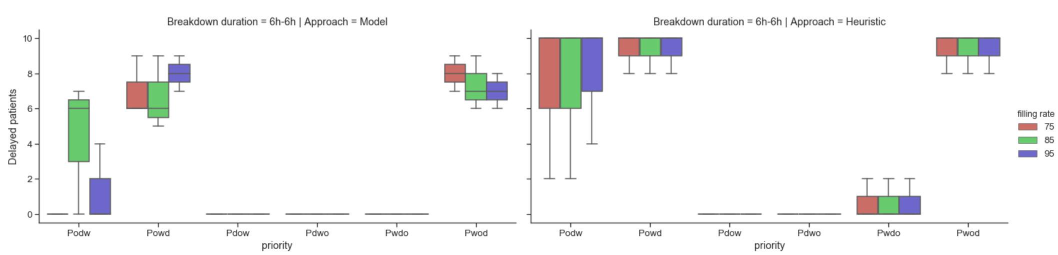

Figure 6.20 Number of delayed patients for 6h machine breakdown

CHAPITRE 7 DISCUSSION GÉNÉRALE

Cette thèse contient des contributions importantes en santé. Nous proposons trois problèmes: la planification des rendez-vous des patients en utilisant un modèle de prédiction, la planification des rendez-vous des patients et des horaires des technologues, et la replanification des rendez-vous des patientes suite à une panne des machines.

7.1 Synthèse des travaux

Dans le chapitre 4, nous présentons une étude basée sur les données réelles du CICL afin d'améliorer le système de prise des rendez-vous au centre de radiothérapie. En premier lieu, nous élaborons un modèle performant de prédiction des durées des traitements. En nous basant sur les résultats de prédiction, nous reconstruisons les grilles des rendez-vous dans le but de maximiser le nombre de patients traités par journée.

Dans le chapitre 5, nous évaluons l'impact de la combinaison de la confection des grilles des patients et des horaires des technologues. Nous développons un modèle mathématique en deux versions. La première version commence par la planification des horaires des technologues, par la suite, les grilles des rendez-vous est réalisée; or, la deuxième version intègre les deux simultanément. Les deux versions du modèle sont appliquées dans le cas du département d'IRM au CHUM. Plusieurs scénarios sont évalués en changeant les règles de travail des technologues et les méthodes de construction de leurs plannings.

Dans le chapitre 6, nous développons un outil d'aide à la décision qui permet d'évaluer l'impact des décisions de replanification sur la performance. L'approche présentée est basée sur un modèle d'optimisation et une heuristique. Nous penons en considération toutes les séquences de priorité des décisions de replanification qui sont : retarder un patient, surréserver un patient, utiliser le temps supplémentaire pour céduler un patient. La solution proposée vise à minimiser le nombre de patients retardés, le temps d'attente et le temps supplémentaire.

7.2 Limitations des solutions proposées

Malgré les bons résultats récoltés dans cette thèse, il existe toujours des limitations qui empêchent la recherche de réaliser plus de progrès.

La nature des attributs utilisés dans la prédiction du temps de traitement est une des limitations du premier objectif de cette thèse. La prédiction est faite en se basant sur les caractéristiques du traitement des patients, mais nous n'avons pas de l'information concernant les attributs des patients tels que l'âge, le poids, etc. L'accès à plus de détails pourrait améliorer davantage la qualité du modèle de prédiction.

Le manque de la documentation liée à la panne des machines en radiothérapie est une des limitations du troisième article de cette thèse. L'accès à cette information permettra de faire une bonne analyse statistique et de proposer des solutions plus appropriées à notre étude de cas.

7.3 Améliorations futures

L'importance et la réalité des sujets traités dans la présente thèse nous poussent à chercher des potentiels d'amélioration des systèmes qui existent actuellement en santé.

Dans le premier article de la thèse, nous prédisons le temps de service des patients en radiothérapie en nous basant sur les caractéristiques du traitement. Dans le futur, l'introduction d'autres attributs tels que l'âge et le poids permettra d'augmenter l'efficience de l'approche proposée.

Dans cette thèse, nous présentons une planification intégrée des rendez-vous des patients et des horaires des technologues en analysant l'impact du nombre des technologues alloués sur le nombre de patients traités. Cette approche est appliquée au cas du centre d'imagerie. Dans le futur, nous pouvons l'utiliser pour d'autres contextes dont le système de planification des rendez-vous est basé sur les grilles, et le nombre du personnel soignant influence le temps de service.

Dans le troisième article présenté dans cette thèse, nous proposons une amélioration de l'heuristique de replanification des patients en radiothérapie à l'aide de la recherche locale.

Cette méthode a prouvé son efficacité. Dans le futur, nous pouvons utiliser des techniques plus complexes pour améliorer davantage l'heuristique, ce qui permettra d'avoir un outil performant et facile à implémenter.

CHAPITRE 8 CONCLUSION ET RECOMMANDATIONS

L'exploitation des techniques d'apprentissage machine et de la recherche opérationnelle a prouvé sa performance dans l'amélioration du système de santé. En effet, la planification efficiente des rendez-vous des patients et l'optimisation des horaires des technologues ont permis de favoriser l'accès de patients aux soins, la bonne allocation des ressources humaines et matérielles, ainsi que la minimisation des coûts d'attente et du temps supplémentaire. Dans cette thèse, nous avons traité trois sujets de recherche sous des collaborations avec le CICL et le CHUM. Nous avons intégré toutes les contraintes réelles des problèmes proposées afin d'étudier la complexité actuelle confrontée par les gestionnaires.

Nous avons considéré les décisions de planification à travers deux niveaux hiérarchiques : tactique et opérationnelle. Le premier objectif de la thèse combine les deux niveaux d'une manière séquentielle. En premier lieu, nous commençons par le plus haut niveau, nous avons redéfini la longueur des créneaux horaires dans les grilles des rendez-vous; par la suite, les grilles ont été comparées au niveau opérationnel en appliquant des règles de gestion et de séquences des patients. Le deuxième objectif s'intègre au niveau tactique, nous générons mensuellement les grilles des rendez-vous des patients et les horaires des technologues en évaluant les règles de travail des technologues et les méthodes de construction des plannings. Le troisième objectif attaque le problème d'occurrence des évènements imprévus tels que la panne des machines au niveau opérationnel, nous replanifions la planification des rendez-vous des patients en radiothérapie en minimisant le nombre des patients retardés, le temps d'attente, et le temps supplémentaire.

Le grand défi des projets de recherche appliqués en santé est de proposer des solutions simples, efficaces, et qui peuvent être facilement analysées par le personnel de santé qui n'est pas expert en mathématiques. Dans les trois problèmes traités dans cette thèse, nous avons guidé les gestionnaires vers une bonne exploitation des solutions proposées en leur montrant le meilleur plan pour tirer profit des analyses des résultats trouvés.

RÉFÉRENCES

AHMADI-JAVID, A., JALALI, Z. et KLASSEN, K. J. (2017). Outpatient appointment systems in healthcare: A review of optimization studies. *European Journal of Operational Research*, 258(1):3–34.

AKBARZADEH, B., MOSLEHI, G., REISI-NAFCHI, M. et MAENHOUT, B. (2019). The replanning and scheduling of surgical cases in the operating room department after block release time with resource rescheduling. *European Journal of Operational Research*, 278(2): 596–614.

ALAEDDINI, A., YANG, K., REDDY, C. et YU, S. (2011). A probabilistic model for predicting the probability of no-show in hospital appointments. *Health Care Management Science*, 14(2):146–157.

ALFARES, H. K. (2004). Survey, categorization, and comparison of recent tour scheduling literature. *Annals of Operations Research*, 127(1-4):145–175.

ALKHATIB, O. et ALAHMAR, A. (2021). A literature review on length of stay prediction for stroke patients using machine learning and statistical approaches. *arXiv preprint arXiv:2201.00005*.

ASSOCIATION CANADIENNE DES RADIOLOGISTES (2020). Radiology resilience now and beyond: Report from the canadian radiology resilience taskforce.

BACCHUS, B. et MACKENZIE, M. (2019). Waiting your turn: Wait times for health care in canada, 2019 report. *Fraser Institute,* *https://www.fraserinstitute.org/sites/default/files/waiting-your-turn-2019-execsum.pdf*.

BAILEY, N. T. (1952). A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times. *Journal of the Royal Statistical Society: Series B (Methodological)*, 14(2):185–199.

BAKER, K. R. (1976). Workforce allocation in cyclical scheduling problems: A survey. *Journal of the Operational Research Society*, 27(1):155–167.

BAKKER, M. et TSUI, K.-L. (2017). Dynamic resource allocation for efficient patient scheduling: A data-driven approach. *Journal of Systems Science and Systems Engineering*, 26(4):448–462.

BENTAYEB, D., LAHRICHI, N. et ROUSSEAU, L.-M. (2019). Patient scheduling based on a service-time prediction model: a data-driven study for a radiotherapy center. *Health care management science*, 22(4):768–782.

BILLINGS, J., DIXON, J., MIJANOVICH, T. et WENNBERG, D. (2006). Case finding for patients at risk of readmission to hospital: Development of algorithm to identify high risk patients. *BMJ*, 333(7563):327.

BOONMEE, C., PISUTHA-ARNOND, N., CHATTINNAWAT, W., MUANGWONG, P., NOBNOP, W. et CHITAPANARUX, I. (2021). Decision support system for radiotherapy patient scheduling: Thai cancer center case study. *In 2021 5th International Conference on Medical and Health Informatics*, pages 168–175.

BRAGA, P., PORTELA, F., SANTOS, M. F. et RUA, F. (2014). Data mining models to predict patient's readmission in intensive care units. *In ICAART 2014-Proceedings of the 6th International Conference on Agents and Artificial Intelligence*.

BRAUNE, R., GUTJAHR, W. J. et VOGL, P. (2021). Stochastic radiotherapy appointment scheduling. *Central European Journal of Operations Research*, pages 1–39.

BRUNNER, J. O., BARD, J. F. et KOLISCH, R. (2009). Flexible shift scheduling of physicians. *Health care management science*, 12(3):285–305.

BRUNNER, J. O., BARD, J. F. et KOLISCH, R. (2010). Midterm scheduling of physicians with flexible shifts using branch and price. *IIE Transactions*, 43(2):84–109.

BURKE, E. K., DE CAUSMAECKER, P., BERGHE, G. V. et VAN LANDEGHEM, H. (2004). The state of the art of nurse rostering. *Journal of scheduling*, 7(6):441–499.

Castro, E. et Petrovic, S. (2012). Combined mathematical programming and heuristics for a radiotherapy pre-treatment scheduling problem. *Journal of Scheduling*, 15(3):333–346.

Cayirli, T. et Veral, E. (2003). Outpatient scheduling in health care: A review of literature. *Production and Operations Management*, 12(4):519–549.

Cayirli, T., Veral, E. et Rosen, H. (2006). Designing appointment scheduling systems for ambulatory care services. *Health Care Management Science*, 9(1):47–58.

Chen, P.-S., Chen, G. Y.-H., Liu, L.-W., Zheng, C.-P. et Huang, W.-T. (2022). Using simulation optimization to solve patient appointment scheduling and examination room assignment problems for patients undergoing ultrasound examination. *In Healthcare*, volume 10, page 164. MDPI.

Chen, P.-S., Lin, Y.-J. et Peng, N.-C. (2016). A two-stage method to determine the allocation and scheduling of medical staff in uncertain environments. *Computers & Industrial Engineering*, 99:174–188.

Choy, G., Khalilzadeh, O., Michalski, M., Do, S., Samir, A. E., Pianykh, O. S., Geis, J. R., Pandharipande, P. V., Brink, J. A. et Dreyer, K. J. (2018). Current applications and future impact of machine learning in radiology. *Radiology*, 288(2):318.

Clark, A., Moule, P., Topping, A. et Serpell, M. (2015). Rescheduling nursing shifts: scoping the challenge and examining the potential of mathematical model based tools. *Journal of Nursing Management*, 23(4):411–420.

Condotta, A. et Shakhlevich, N. (2014). Scheduling patient appointments via multilevel template: A case study in chemotherapy. *Operations Research for Health Care*, 3(3):129–144.

Conforti, D., Guerriero, F. et Guido, R. (2008). Optimization models for radiotherapy patient scheduling. *4or*, 6(3):263–278.

Conforti, D., Guerriero, F. et Guido, R. (2010). Non-block scheduling with priority for radiotherapy treatments. *European Journal of Operational Research*, 201(1):289–296.

Conforti, D., Guerriero, F., Guido, R. et Veltri, M. (2011). An optimal decision-making approach for the management of radiotherapy patients. *OR Spectrum*, 33(1):123–148.

Dantas, L. F., Hamacher, S., Cyrino Oliveira, F. L., Barbosa, S. D. et Viegas, F. (2019). Predicting patient no-show behavior: a study in a bariatric clinic. *Obesity surgery*, 29(1):40–47.

Davarian, F. et Behnamian, J. (2022). Robust finite-horizon scheduling/rescheduling of operating rooms with elective and emergency surgeries under resource constraints. *Journal of Scheduling*, 25(6):625–641.

Denton, B., Viapiano, J. et Vogl, A. (2007). Optimization of surgery sequencing and scheduling decisions under uncertainty. *Health Care Management Science*, 10(1):13–24.

Donnan, P. T., Dorward, D. W., Mutch, B. et Morris, A. D. (2008). Development and validation of a model for predicting emergency admissions over the next year (PEONY): A UK historical cohort study. *Archives of Internal Medicine*, 168(13):1416–1422.

Dunning, I., Huchette, J. et Lubin, M. (2017). Jump: A modeling language for mathematical optimization. *SIAM Review*, 59(2):295–320.

Elkin, E. B., Snow, J. G., Leoce, N. M., Atoria, C. L. et Schrag, D. (2012). Mammography capacity and appointment wait times: barriers to breast cancer screening. *Cancer Causes & Control*, 23(1):45–50.

Erekat, A., Servis, G., Madathil, S. C. et Khasawneh, M. T. (2020). Efficient operating room planning using an ensemble learning approach to predict surgery cancellations. *IISE Transactions on Healthcare Systems Engineering*, 10(1):18–32.

Erhard, M., Schoenfelder, J., Fügener, A. et Brunner, J. O. (2018). State of the art in physician scheduling. *European Journal of Operational Research*, 265(1):1–18.

Ernst, A. T., Jiang, H., Krishnamoorthy, M. et Sier, D. (2004). Staff scheduling and rostering: A review of applications, methods and models. *European journal of operational research*, 153(1):3–27.

Fan, G., Deng, Z., Ye, Q. et Wang, B. (2021). Machine learning-based prediction models for patients no-show in online outpatient appointments. *Data Science and Management*, 2:45–52.

Gilles, S., Ania, K., Casey, H. et Heidi, S. (2022). Rétablir un accès à l'imagerie médicale dans des délais acceptables au canada : de nouveaux investissements dans la radiologie sont nécessaires.

Glover IV, M., Daye, D., Khalilzadeh, O., Pianykh, O., Rosenthal, D. I., Brink, J. A. et Flores, E. J. (2017). Socioeconomic and demographic predictors of missed opportunities to provide advanced imaging services. *Journal of the American College of Radiology*, 14(11):1403–1411.

Glowacka, K. J., Henry, R. M. et May, J. H. (2009). A hybrid data mining/simulation approach for modelling outpatient no-shows in clinic scheduling. *Journal of the Operational Research Society*, 60(8):1056–1068.

Gocgun, Y. (2018). Simulation-based approximate policy iteration for dynamic patient scheduling for radiation therapy. *Health care management science*, 21(3):317–325.

Golmohammadi, D. (2021). A decision-making tool based on historical data for service time prediction in outpatient scheduling. *International Journal of Medical Informatics*, 156:104591.

Golmohammadi, D. et Radnia, N. (2016). Prediction modeling and pattern recognition for patient readmission. *International Journal of Production Economics*, 171:151–161.

Gross, C. N., Fügener, A. et Brunner, J. O. (2018). Online rescheduling of physicians in hospitals. *Flexible Services and Manufacturing Journal*, 30(1):296–328.

Gunasekaran, G. H., Hassali, M. A. B. A., Sabri, W. M. A. B. W. et Rahman, M. T. B. (2020). Impact of chemotherapy schedule modification on breast cancer patients: a single-centre retrospective study. *International journal of clinical pharmacy*, 42(2):642–651.

Gupta, D. (2007). Surgical suites' operations management. *Production and Operations Management*, 16(6):689–700.

GUPTA, D. et DENTON, B. (2008). Appointment scheduling in health care: Challenges and opportunities. *IIE Transactions*, 40(9):800–819.

HADID, M., ELOMRI, A., EL MEKKAWY, T., KERBACHE, L., EL OMRI, A., EL OMRI, H., TAHA, R. Y., HAMAD, A. A. et AL THANI, M. H. J. (2022). Bibliometric analysis of cancer care operations management: current status, developments, and future directions. *Health Care Management Science*, 25(1):166–185.

HARRIS, S. L., MAY, J. H. et VARGAS, L. G. (2016). Predictive analytics model for health-care planning and scheduling. *European Journal of Operational Research*, 253(1):121–131.

HÉBERT, G., SULLY, J.-L. et NGUYEN, M. (2017). *Allocation des ressources pour la santé et les services sociaux au Québec.* Institut de recherche et d'informations socio-économiques.

HOOSHANGI-TABRIZI, P., CONTRERAS, I., BHUIYAN, N. et BATIST, G. (2020). Improving patient-care services at an oncology clinic using a flexible and adaptive scheduling procedure. *Expert Systems with Applications*, 150:113267.

HUANG, Y. et HANAUER, D. A. (2014). Patient no-show predictive model development using multiple data sources for an effective overbooking approach. *Applied Clinical Informatics*, 5(3):836–860.

HUANG, Y. et VERDUZCO, S. (2015). Appointment template redesign in a women's health clinic using clinical constraints to improve service quality and efficiency. *Applied clinical informatics*, 6(02):271–287.

HUANG, Y.-L. et BACH, S. M. (2016). Appointment lead time policy development to improve patient access to care. *Applied Clinical Informatics*, 7(4):954–968.

HUANG, Y.-L., BANSAL, A., BERG, B. P., TOMMASO, C. P. et LAUGHLIN, R. S. (2022). Coordination of intraoperative neurophysiologic monitoring technologist and surgery schedules. *Journal of Medical Systems*, 46(10):1–11.

HUANG, Y.-L. et MARCAK, J. (2013). Radiology scheduling with consideration of patient characteristics to improve patient access to care and medical resource utilization. *Health Systems*, 2(2):93–102.

Institut canadien d'information sur la santé (2021a). Tendances des dépenses nationales de santé, 2021 — analyse éclair.

Institut canadien d'information sur la santé (2021b). Votre système de santé.

Institut canadien d'information sur la santé (2022). Temps d'attente pour la radiothérapie (percentiles).

Jain, D. et Singh, V. (2016). Utilization of data mining classification approach for disease prediction: A survey. *International Journal Education and Management Engineering*, 6:45–52.

Kapamara, T., Sheibani, K., Haas, O. C., Reeves, C. R. et Petrovic, D. (2006). A review of scheduling problems in radiotherapy. *In Proceedings of the Eighteenth International Conference on Systems Engineering (ICSE2006), Coventry University, UK*, pages 201–207.

Kim, S.-H., Whitt, W. et Cha, W. C. (2018). A data-driven model of an appointment-generated arrival process at an outpatient clinic. *INFORMS Journal on Computing*, 30(1): 181–199.

Koh, H. C., Tan, G. *et al.* (2011). Data mining applications in healthcare. *Journal of Healthcare Information Management*, 19(2):64–72.

Kolisch, R. et Sickinger, S. (2008). Providing radiology health care services to stochastic demand of different customer classes. *OR spectrum*, 30(2):375–395.

Kuo, Y.-H., Chan, N. B., Leung, J. M., Meng, H., So, A. M.-C., Tsoi, K. K. et Graham, C. A. (2020). An integrated approach of machine learning and systems thinking for waiting time prediction in an emergency department. *International journal of medical informatics*, 139:104143.

Lagoe, R. J., Noetscher, C. M. et Murphy, M. P. (2001). Hospital readmission: Predicting the risk. *Journal of Nursing Care Quality*, 15(4):69–83.

Larose, D. T. et Larose, C. D. (2014). *Discovering Knowledge in Data: An Introduction to Data Mining.* John Wiley & Sons.

LEGRAIN, A. (2015). *Optimisation stochastique de problèmes d'ordonnancement en santé.* Thèse de doctorat, École Polytechnique de Montréal.

LEGRAIN, A., FORTIN, M.-A., LAHRICHI, N. et ROUSSEAU, L.-M. (2015). Online stochastic optimization of radiotherapy patient scheduling. *Health care management science*, 18(2): 110–123.

LI, Z. et IERAPETRITOU, M. (2008). Process scheduling under uncertainty: Review and challenges. *Computers & Chemical Engineering*, 32(4-5):715–727.

LONG, Z., WEN, X., LAN, M. et YANG, Y. (2022). Nursing rescheduling problem with multiple rescheduling methods under uncertainty. *Complex & Intelligent Systems*, 8(6): 4557–4569.

LOTFI, V. et TORRES, E. (2014). Improving an outpatient clinic utilization using decision analysis-based patient scheduling. *Socio-Economic Planning Sciences*, 48(2):115–126.

MAGESHWARI, G. et KANAGA, E. G. M. (2012). Literature review on patient scheduling techniques. *International Journal on Computer Science and Engineering*, 4(3):397.

MANDELBAUM, A., MOMČILOVIĆ, P., TRICHAKIS, N., KADISH, S., LEIB, R. et BUNNELL, C. A. (2020). Data-driven appointment-scheduling under uncertainty: The case of an infusion unit in a cancer center. *Management Science*, 66(1):243–270.

MORADI, S., NAJAFI, M., MESGARI, S. et ZOLFAGHARINIA, H. (2022). The utilization of patients' information to improve the performance of radiotherapy centers: A data-driven approach. *Computers & Industrial Engineering*, 172:108547.

OGULATA, S. N., KOYUNCU, M. et KARAKAS, E. (2008). Personnel and patient scheduling in the high demanded hospital services: a case study in the physiotherapy service. *Journal of medical systems*, 32(3):221–228.

OUELHADJ, D. et PETROVIC, S. (2009). A survey of dynamic scheduling in manufacturing systems. *Journal of scheduling*, 12(4):417–431.

Patrick, J. et Puterman, M. L. (2007). Improving resource utilization for diagnostic services through flexible inpatient scheduling: A method for improving resource utilization. *Journal of the Operational Research Society*, 58(2):235–245.

Patrick, J., Puterman, M. L. et Queyranne, M. (2008). Dynamic multipriority patient scheduling for a diagnostic resource. *Operations research*, 56(6):1507–1525.

Pereira, S., Torres, L., Portela, F., Santos, M. F., Machado, J. et Abelha, A. (2016). Predicting triage waiting time in maternity emergency care by means of data mining. *In New Advances in Information Systems and Technologies*, pages 579–588. Springer.

Pérez, E., Ntaimo, L., Malavé, C. O., Bailey, C. et McCormack, P. (2013). Stochastic online appointment scheduling of multi-step sequential procedures in nuclear medicine. *Health care management science*, 16(4):281–299.

Pérez, E., Ntaimo, L., Wilhelm, W. E., Bailey, C. et McCormack, P. (2011). Patient and resource scheduling of multi-step medical procedures in nuclear medicine. *IIE Transactions on Healthcare Systems Engineering*, 1(3):168–184.

Petrovic, D., Morshed, M. et Petrovic, S. (2009). Genetic algorithm based scheduling of radiotherapy treatments for cancer patients. *In Conference on Artificial Intelligence in Medicine in Europe*, pages 101–105. Springer.

Petrovic, S. et Castro, E. (2011). A genetic algorithm for radiotherapy pre-treatment scheduling. *In European conference on the applications of evolutionary computation*, pages 454–463. Springer.

Petrovic, S. et Leite-Rocha, P. (2008). Constructive and grasp approaches to radiotherapy treatment scheduling. *In Advances in Electrical and Electronics Engineering-IAENG Special Edition of the World Congress on Engineering and Computer Science 2008*, pages 192–200. IEEE.

Petrovic, S., Leung, W., Song, X. et Sundar, S. (2006). Algorithms for radiotherapy treatment booking. *In 25th Workshop of the UK planning and scheduling special interest group*, pages 105–112. Citeseer.

SAURE, A., PATRICK, J., TYLDESLEY, S. et PUTERMAN, M. L. (2012). Dynamic multi-appointment patient scheduling for radiation therapy. *European Journal of Operational Research*, 223(2):573–584.

SIMSEK, S., TIAHRT, T. et DAG, A. (2020). Stratifying no-show patients into multiple risk groups via a holistic data analytics-based framework. *Decision Support Systems*, 132:113269.

SOCIÉTÉ CANADIENNE DU CANCER (2017). Statistiques canadiennes sur le cancer.

SOCIÉTÉ CANADIENNE DU CANCER (2021). Statistiques canadiennes sur le cancer.

SRINIVAS, S. et RAVINDRAN, A. R. (2018). Optimizing outpatient appointment system using machine learning algorithms and scheduling rules: a prescriptive analytics framework. *Expert Systems with Applications*, 102:245–261.

SRINIVAS, S. et SALAH, H. (2021). Consultation length and no-show prediction for improving appointment scheduling efficiency at a cardiology clinic: a data analytics approach. *International Journal of Medical Informatics*, 145:104290.

STOLLETZ, R. et BRUNNER, J. O. (2012). Fair optimization of fortnightly physician schedules with flexible shifts. *European Journal of Operational Research*, 219(3):622–629.

STUART, K. et KOZAN, E. (2012). Reactive scheduling model for the operating theatre. *Flexible Services and Manufacturing Journal*, 24(4):400–421.

TASGETIREN, M. F., LIANG, Y.-C., SEVKLI, M. et GENCYILMAZ, G. (2007). A particle swarm optimization algorithm for makespan and total flowtime minimization in the permutation flowshop sequencing problem. *European journal of operational research*, 177(3):1930–1947.

TOPUZ, K., UNER, H., OZTEKIN, A. et YILDIRIM, M. B. (2018). Predicting pediatric clinic no-shows: a decision analytic framework using elastic net and bayesian belief network. *Annals of Operations Research*, 263(1):479–499.

UHLMANN, I. R. et FRAZZON, E. M. (2018). Production rescheduling review: Opportunities for industrial integration and practical applications. *Journal of manufacturing systems*, 49:186–193.

VALI-SIAR, M. M., GHOLAMI, S. et RAMEZANIAN, R. (2018). Multi-period and multi-resource operating room scheduling under uncertainty: A case study. *Computers & Industrial Engineering*, 126:549–568.

Van den BERGH, J., BELIËN, J., DE BRUECKER, P., DEMEULEMEESTER, E. et DE BOECK, L. (2013). Personnel scheduling: A literature review. *European journal of operational research*, 226(3):367–385.

VAN NYNATTEN, L. et GERSHON, A. (2017). Radiology wait times. *University of Western Ontario Medical Journal*, 86(2):65–66.

VAN SAMBEEK, J. R., JOUSTRA, P. E., DAS, S. F., BAKKER, P. J. et MAAS, M. (2011). Reducing mri access times by tackling the appointment-scheduling strategy. *BMJ Qual Saf*, 20(12):1075–1080.

van WALRAVEN, C., WONG, J., HAWKEN, S. et FORSTER, A. J. (2012). Comparing methods to calculate hospital-specific rates of early death or urgent readmission. *Canadian Medical Association Journal*, 184(15):E810–E817.

VIEIRA, B., DEMIRTAS, D., van de KAMER, J. B., HANS, E. W., ROUSSEAU, L.-M., LAHRICHI, N. et van HARTEN, W. H. (2020). Radiotherapy treatment scheduling considering time window preferences. *Health care management science*, 23(4):520–534.

VIEIRA, B., DEMIRTAS, D., van de KAMER, J. B., HANS, E. W. et van HARTEN, W. (2018). A mathematical programming model for optimizing the staff allocation in radiotherapy under uncertain demand. *European journal of operational research*, 270(2):709–722.

VIEIRA, B., HANS, E. W., van VLIET-VROEGINDEWEIJ, C., VAN DE KAMER, J. et VAN HARTEN, W. (2016). Operations research for resource planning and-use in radiotherapy: a literature review. *BMC medical informatics and decision making*, 16(1):1–11.

Vieira, G. E., Herrmann, J. W. et Lin, E. (2003). Rescheduling manufacturing systems: a framework of strategies, policies, and methods. *Journal of scheduling*, 6(1):39–62.

Walter, S. (1973). A comparison of appointment schedules in a hospital radiology department. *British journal of preventive & social medicine*, 27(3):160.

Wang, B., Han, X., Zhang, X. et Zhang, S. (2015). Predictive-reactive scheduling for single surgical suite subject to random emergency surgery. *Journal of Combinatorial Optimization*, 30(4):949–966.

Wang, H., Wang, W., Sun, H., Cui, Z., Rahnamayan, S. et Zeng, S. (2017). A new cuckoo search algorithm with hybrid strategies for flow shop scheduling problems. *Soft Computing*, 21(15):4297–4307.

Wang, X. et Tang, L. (2012). A discrete particle swarm optimization algorithm with self-adaptive diversity control for the permutation flowshop problem with blocking. *Applied Soft Computing*, 12(2):652–662.

Yuura, H., Miyamoto, T. et Hidaka, K. (2017). An integer programming model for radiographer scheduling considering skills and training. *In 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pages 889–893. IEEE.