European Journal of Operational Research xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Scheduling the hospital-wide flow of elective patients

Daniel Gartner *, Rainer Kolisch

TUM School of Management, Technische Universität München, Germany

ARTICLE INFO

Article history:
Received 9 August 2012
Accepted 19 August 2013
Available online xxxx

Keywords:
Clinical pathways
Diagnosis-related groups
Patient flow management

ABSTRACT

In this paper, we address the problem of planning the patient flow in hospitals subject to scarce medical resources with the objective of maximizing the contribution margin. We assume that we can classify a large enough percentage of elective patients according to their diagnosis-related group (DRG) and clinical pathway. The clinical pathway defines the procedures (such as different types of diagnostic activities and surgery) as well as the sequence in which they have to be applied to the patient. The decision is then on which day each procedure of each patient's clinical pathway should be done, taking into account the sequence of procedures as well as scarce clinical resources, such that the contribution margin of all patients is maximized. We develop two mixed-integer programs (MIP) for this problem which are embedded in a static and a rolling horizon planning approach. Computational results on real-world data show that employing the MIPs leads to a significant improvement of the contribution margin compared to the contribution margin obtained by employing the planning approach currently practiced. Furthermore, we show that the time between admission and surgery is significantly reduced by applying our models.

1. Introduction

For many years, cost reimbursement has been the standard payment scheme for hospitals. In this scheme, a hospital receives the total cost for treating the patient which is calculated by multiplying the patient's length of stay with a (ward specific) daily rate and adding the costs for clinical procedures applied, such as diagnosis and surgery. Cost reimbursement does not provide an incentive for hospitals to operate efficiently. As a consequence, hospital costs as the largest part of total health care costs (approximately 31% in the US, see Lim, Mobasher, Kardar, & Cote (2011, chap. 3)), have increased sharply. In an effort to limit hospital costs and to create an incentive for hospitals to operate more efficiently, many countries have introduced payment schemes that are based on diagnosis-related groups (DRGs). In these schemes, patients are classified into DRGs with homogeneous clinical characteristics and resources required during treatment within each group, while between groups, the patients' clinical characteristics and therefore costs are different. Hospitals receive payments based on the DRG instead of the applied procedures and the length of stay. The reimbursement the hospital receives for treating a patient with a specific DRG equals the average cost which accrued in a representative sample of hospitals in the year before last. For each country, there is an institution, such as the German institute for the reimbursement in hospitals INEK (see Schreyögg, Tiemann, & Busse (2006)), responsible for generating the sample, collecting, processing and distributing the cost data. With DRG-based payment schemes in place, in order to be profitable a hospital has to operate such that its patient specific costs are lower than the average costs of the hospitals in the sample. However, the cost structure in hospitals, as in many service industries, is such that the predominant share of the operational costs is fixed. Hence, costs can only be marginally reduced by operational decisions. Instead, hospitals have to use existing resources such that revenue or the total contribution margin is maximized. This approach is well-known in the service industry (see Kimes (1989) for an example in the service industry in general as well as Schütz & Kolisch (2012) for an example in the health care industry in particular).

In this paper, we propose two models to plan the patient flow in hospitals subject to scarce medical resources with the objective to maximize the contribution margin. We assume that we can classify a large enough percentage of elective patients according to DRG and clinical pathway. The clinical pathway defines the procedures (such as different types of diagnostic activities and surgery) as well as the sequence in which they have to be applied to the patient. The decision is then on which day each procedure of each patient's clinical pathway should be done, taking into account the sequence of procedures as well as scarce clinical resources, such that the contribution margin is maximized. We consider a particular variant of the DRG-based payment scheme which is in use in Australia

* Corresponding author. Tel.: +49 8928925161; fax: +49 8928925166.
E-mail addresses: daniel.gartner@tum.de (D. Gartner), rainer.kolisch@tum.de (R. Kolisch).

0377-2217/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2013.08.026

(see Duckett (1998)), Switzerland (see Zaugg, Gattiker, Moneta, & Reay (2012)), France and Germany (see Schreyögg, Tiemann et al. (2006)). In this variant, payment is not only dependent on DRG but also, to some extent, on length of stay. More precisely, the fixed reimbursement is reduced by some fixed amount per day if the length of stay (LOS) falls below a lower threshold value, referred to in the literature as low LOS trim point. In contrast, the fixed reimbursement is increased by some fixed amount per day if the medically necessary length of stay exceeds a high threshold value, denoted by high LOS trim point. By this, the payment scheme takes outliers into account (see, for example, Schreyögg, Tiemann et al. (2006)). The concept of trim points is described in Section 3; for more information about DRG trim points, see Busse, Geissler, Quentin, and Wiley (2011). An example of a cost, revenue and contribution margin function is given in Section 3.1, while a detailed example of two patients and their corresponding contribution margin for two specific DRGs is provided in Section 3.3. The payment scheme with fixed payment not coupled to the length of stay is a special case of the one that is treated in our study. Hence, our models can be applied to both variants.

Within the health care planning matrix of Hans, van Houdenhoven, and Hulshof (2011) our planning approach can be classified as offline operational resource capacity planning. It comprises the detailed coordination of the activities regarding current (elective) demand (see Hulshof, Kortbeek, Boucherie, Hans, & Bakker (2012)). Our models thus use planning decisions which have been made at higher planning levels such as strategic decisions on case mix, type and capacity of clinical resources as well as tactical decisions on the master surgery schedule. Then, our models generate decisions on the admission day of patients, the day-based assignment of clinical activities to hospital resources and the discharge day of patients. These results are in turn input for decisions to be made on subordinate planning levels such as bed assignment in wards, sequencing of diagnostic activities and surgical cases within a day. Contrary to most planning approaches for the offline operational level proposed so far, our model considers not only one resource such as a diagnostic device or the operating theater, but all scarce clinical resources. Correspondingly, we do not consider a single type of activity within the clinical pathway, such as the surgery, but all activities. We manage to do so by applying the right amount of aggregation in our models with a period length of one day. Since the capacity demand of a single activity in a clinical pathway is usually significantly smaller than a day, we can use expected values instead of distribution functions for most data. However, for some data such as recovery times of patients with a length of multiple days and arrival times of patients we explicitly take the stochastic nature into account by embedding our models in a rolling horizon approach with Bayesian recovery time prediction. Our computational results using real-world data demonstrate that our approaches lead to an improvement of approximately 5% in the contribution margin when compared to the planning approach currently employed in the hospital from which we obtained our data. Also, we show that the time span between admission and surgery is reduced by about 1.5 days and that the length of stay is reduced by about 2 days.

The remainder of this paper is structured as follows. In Section 2, we provide a survey of the relevant literature and set apart our approach. Section 3 presents the mathematical models and illustrates them by means of examples. In Section 4, we test our models on real-world data from a midsize hospital. First, we provide results for the two static deterministic models of Section 3. Then, we show how the models are embedded within a rolling horizon approach using Bayesian recovery time prediction in order to take into account uncertain information. We provide a computational comparison of the static deterministic and the dynamic stochastic planning, benchmarking both against the solution approach currently employed by the hospital. Finally, we analyze the impact of our approaches on some key figures such as time span between admission and surgery as well as length of stay. Section 5 closes the paper with conclusions.

2. Literature review

There is a rich literature on offline operational capacity planning in hospitals. Table 1 provides an overview of offline operational capacity planning problems for surgical, inpatient and residential care services, employing the taxonomy of Hulshof et al. (2012). We selected articles based on that taxonomy and restricted our search to the application of mathematical programming as a solution technique. Moreover, we excluded staff-to-shift assignment papers and those that were published before 2000. We have classified each article with respect to the use of scarce clinical resources (single vs. multiple) and the consideration of risk (deterministic vs. stochastic). The table reveals that the majority of the papers consider a single resource such as the operating theater (for a survey see Cardoen, Demeulemeester, & Beliën (2010) and Guerriero & Guido (2011)) or wards (Demeester, Souffriau, De Causmaecker, & Vanden Berghe (2009)). The planning of diagnostic services is addressed, for example, by Patrick, Puterman, and Queyranne (2008) who, in contrast to our work, employ Markov decision processes as a solution approach. Since our paper addresses all scarce hospital resources instead of a single one, we do not present these approaches in detail. A smaller number of papers take into account multiple resources, see Vanberkel, Boucherie, Hans, Hurink, and Litvak (2010) for an overview. The majority of the papers which consider multiple resources deal with the operating room in combination with its upstream or downstream resources. Pham and Klinkert (2008) propose a mixed-integer program (MIP) for surgical case scheduling considering scarce resources at the preoperative, perioperative and postoperative stages, solved with the standard solver CPLEX. Augusto, Xie, and Perdomo (2010) consider the problem of scheduling a fixed number of elective cases subject to scarce resources in the operating theater, the post-anesthesia care unit and the patient transportation unit. The problem is modeled as a MIP and solved by using Lagrangian relaxation. Besides the scheduling decision, the model is used to explore the benefit of letting patients recover in the operating room when there is no capacity in the post-anesthesia care unit. Conforti, Guerriero, Guido, Cerinic, and Conforti (2011) address the so-called "week hospital problem" in which a decision is made if and when elective patients on a waiting list are admitted to the hospital and when the clinical activities of the admitted patients are performed. By definition, the week hospital problem ensures that all admitted patients are discharged in the week they have been admitted. The objective is to maximize the sum of the scores of admitted patients. The planning horizon of one week is divided into periods of a half day length. The paper proposes a MIP which is solved with the standard solver CPLEX.

On a tactical planning level, Min and Yih (2010) address the surgery scheduling problem for elective patients taking into account stochastic surgery durations and stochastic capacity of the surgical intensive care unit by employing a stochastic discrete program which is solved with the sample average approximation method. Vissers, Adan, and Bekkers (2005) consider the tactical patient mix optimization problem for a single specialty. Employing a MIP, for each day of the week they decide on the number of patients from different categories to be admitted into the hospital. The scarce clinical resources beds (before surgery and for recovery), the operating theater and the intensive care unit are taken into account. The objective is the minimization of the over- and underutilization of the clinical resources. Apart from the recent