Computers & Operations Research 74 (2016) 78–91

Contents lists available at ScienceDirect

Computers & Operations Research

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

Stochastic programming analysis and solutions to schedule overcrowded operating rooms in China

Guanlian Xiao [a,c], Willem van Jaarsveld [b], Ming Dong [a], Joris van de Klundert [c,*]

[a] Department of Operations Management, Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China
[b] School of Industrial Engineering, Eindhoven University of Technology, Eindhoven, Netherlands
[c] Department of Health Services Management and Organization, Institute of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, Netherlands

ARTICLE INFO

Article history:
Received 18 November 2015
Received in revised form
13 April 2016
Accepted 13 April 2016
Available online 19 April 2016

Keywords:
Operating room scheduling
3-Stage stochastic programming
Sample average approximation
China's hospital reform

ABSTRACT

As a result of the growing demand for health services, China's large city hospitals have become markedly overstretched, resulting in delicate and complex operating room scheduling problems. While the operating rooms are struggling to meet demand, they face idle times because of (human) resources being pulled away for other urgent demands, and cancellations for economic and health reasons. In this research we analyze the resulting stochastic operating room scheduling problems, and the improvements attainable by scheduled cancellations to accommodate the large demand while avoiding the negative consequences of excessive overtime work. We present a three-stage recourse model which formalizes the scheduled cancellations and is anticipative to further uncertainty. We develop a solution method for this three-stage model which relies on the sample average approximation and the L-shaped method. The method exploits the structure of optimal solutions to speed up the optimization. Scheduled cancellations can significantly and substantially improve the operating room schedule when the costs of cancellations are close to the costs of overtime work. Moreover, the proposed methods illustrate how the adverse impact of cancellations (by patients) for economic and health reasons can be largely controlled. The (human) resource unavailability however is shown to cause a more than proportional loss of solution value for the surgery scheduling problems occurring in China's large city hospitals, even when applying the proposed solution techniques, and requires different management measures.

1. Introduction

In the first decade of the present millennium, China's GDP has grown at an average rate of more than 10% [35]. These economic developments have gone hand in hand with social and demographic developments. The urban population grew from 452 million to 721 million [45], the public transportation system improved considerably, and health insurance coverage grew from below 30% around the turn of the millennium to over 95% in 2011 [34]. These changes have driven an enormous growth in demand for health services, and in health expenditures of which 71% are accounted for by hospitals [2]. As a result of these developments, particularly the demand for services at the large (level 3) hospitals in big cities increased [38]. Despite a tenfold growth in government spending on health [28] and a growth in the number of hospitals by more than 40% since the year 2000 [36], the increase in health service capacity has not been able to cope with the rising demand. The level 3 hospitals in big cities have become markedly overstretched [41]. These phenomena are concretely illustrated by the 2013 data provided for the purpose of the analysis presented in this manuscript by Shanghai General Hospital, where the actual average surgical workload exceeded the daily capacity by as much as 20%, and average operating room opening hours are almost 14 h daily.

Because a referral system is lacking, an important part of the increased demand directly reaches the hospitals in the form of ever higher numbers of outpatients, which tend to pull away physicians and other staff from wards and operating rooms. The number of outpatient visits to hospitals has grown from 2.12 billion per year to 3.45 billion per year in the first decade of the new millennium [29]. The increase in outpatient services may cause physicians to be late for operating room shifts or to be called away during operating room shifts, causing idle time at the operating room. (From the complete operating room data for the year 2013, we estimate that idle time at Shanghai General Hospital is around 17%.) In the same decade, the number of inpatient visits in China has more than doubled from 53 million to 133 million annually. Meara et al. [33] recently conservatively estimated the annually

* Corresponding author.
E-mail addresses: xglian@gmail.com (G. Xiao), w.l.v.jaarsveld@tue.nl (W. van Jaarsveld), mdong@sjtu.edu.cn (M. Dong), vandeklundert@bmg.eur.nl (J. van de Klundert).

http://dx.doi.org/10.1016/j.cor.2016.04.017

needed number of surgeries in China at 57 million, of which they considered 27 million to be unmet. The already overstretched operating rooms are therefore likely to face considerable further increases in demand in the coming years. Hence Chinese hospitals face severe operational problems, now and in the coming years.

Our aim is to develop scheduling methods to solve the urgent capacity management problems in China's large city hospitals – which form a priority in the current health system reform – and see how their effectiveness interacts with accompanying operations management measures. As we outline more extensively in the literature review in the next section, current losses of scarce capacity are mostly not due to poor scheduling, but to other causes such as unavailability of scarce (human) resources, and *cancellations* of planned surgeries. Scheduled surgeries may be canceled for a variety of reasons which are beyond the locus of control of operating room management, such as no-show, deteriorating health conditions, and hospital logistics. In anticipation of such *exogenous cancellations*, operating room management may choose to schedule more patients than capacity allows, potentially resulting in capacity problems when cancellations are fewer than expected, or surgeries take longer than expected. The schedulers may subsequently solve the capacity problems by cancelling one or more of the final patients for which surgery was scheduled at the end of the day. Such cancellations may cause dissatisfaction, anxiety and loss of health for the patients, and have led to tense relations between patients and staff. The alternative to further extend overtime hours, on the other hand, is associated with increased risks of complications and medical errors, as well as dissatisfaction among scarce staff ([39] and references therein). The scheduling of operating rooms in the overstretched Chinese hospitals is therefore a stochastic balancing act which is complicated by resource unavailability and exogenous cancellations.

Operating room schedules are typically constructed one or several days in advance. Because of the stochastic nature of surgical services and the related health service processes, schedules are subsequently often adjusted as the day progresses. For operating rooms for elective surgeries, such adjustments are primarily constrained to changes in surgery start times and, when needed, to cancellations of one or more surgeries of the final patients of the day. It is preferable to take such scheduling decisions to cancel one or more of the final patients early, so as to limit the negative effects for patients and staff mentioned above. In practice, such cancellations may also take the form of redirecting patients to another hospital.

The first research objective is now to optimize the operating room schedules. This starts with the optimization of the schedules created one or more days in advance per single elective operating room, henceforth referred to as the first stage problem. Secondly, we consider the optimization of early *scheduled cancellations*, cancellations initiated by the operating room schedulers after an initial part of the daily schedule has been completed (see for instance [39]), referred to as the second stage problem. In particular, we analyze the improvements attainable by introducing a two stage approach (in which the first stage solution takes into account that a second stage follows) over the common practice of a single stage approach which disregards cancellation until the end of the day. The objective will be to balance the benefits from performing surgeries with the costs of overtime work and negative effects of scheduled cancellations. Our modelling of overtime costs reflects the empirical findings that overtime work is increasingly undesirable for patients and staff as the duration lengthens. Moreover, we model resource unavailability and exogenous cancellations as independent stochastic processes and consider surgical durations to be stochastic as well, fitting real life data. As we are interested in the performance improvement possible by adopting a two stage approach, we develop solution methods which solve the problem with and without scheduled cancellations (almost) to optimality. (See Fig. 1 in Section 3.1 for a visualization of the multi-stage model.)

With these solution methods at hand, the second research objective is then to analyze the extent to which scheduling can overcome the difficulties posed by stochastic resource unavailabilities and exogenous cancellations or, alternatively, whether additional operations management measures are required for this purpose. This second research objective is particularly relevant as the literature review below shows that resource unavailability and exogenous cancellations are, to a certain extent, under the control of hospital management. Hence, our results provide insight in how operating room scheduling and hospital management can interact to alleviate China's hospital overcrowding problems.

Section 2 reviews related literature on (surgical) scheduling with cancellations as well as literature on the occurrence and causes of surgical cancellation. Section 3 formally defines the problem and formulates it as a general three-stage model with integer recourse. Section 4 analyzes theoretical model properties which can help to reduce solution times. Section 5 proposes specific solution algorithms for the problem, and finally Section 6 presents numerical results and analysis. The numerical analysis tests the newly developed 3-stage stochastic programming approach by (almost) optimally solving instances derived from 2013 operating room data of Shanghai General Hospital. To this purpose, we fit distributions to the underlying stochastic processes using a complete data set on surgical operations. QQ-plots show that log-normal distributions fit these surgical durations well, and the proposed SAA approach is able to deal with these analytically inconvenient distributions. The computational results provide insight in the benefits attainable by scheduled cancellations for current rates of resource unavailability and exogenous cancellations. Moreover, we consider scenarios in which additional measures are taken to reduce resource unavailability and exogenous cancellations. We conclude by considering practical implications for operating room management and scheduling in China's overcrowded hospitals.

2. Literature review

The phenomena of cancellation, no-show and overbooking have been studied extensively in the operations management literature, mostly originating from revenue management applications in the airline industry [42]. In this setting, no-show refers to passengers not showing up for a flight without giving prior notice, and cancellation to passengers cancelling their booked flights in advance (which is different from the definitions for cancellations provided above). Like it is the case in the surgical scheduling problem we consider, revenue management models typically exploit the expected benefits from overbooking capacity, taking into account that penalties must be paid when the eventual number of patients showing up exceeds capacity. For instance Subramanian et al. [40] consider an application which includes no-show, cancellation and overbooking. While the revenue management problems considered in the airline and hotel industry are essentially different from surgical scheduling, they share general properties and solution approaches. For instance, Karaesmen and Van Ryzin [20] present a two-stage stochastic program to model no-show and overbooking, where cancellations have become known in the second stage (as is partially the case in our model). Lai and Ng [25] propose a stochastic network optimization model for hotel revenue management and use robust optimization techniques to deal with cancellations, no-show and over-booking of hotel guests. Overbooking has also been introduced in health care, first and