

Patient Flow Analysis in Pain Rehabilitation Care

N. Baer¹², N. Kortbeek¹²³, N. Litvak¹², O. Roukens⁴

¹Stochastic Operations Research, Department of Applied Mathematics, University of Twente,
Drienerlolaan 5, 7500 AE Enschede, The Netherlands;

²Center for Healthcare Operations Improvement and Research (CHOIR),
University of Twente, Drienerlolaan 5, 7500 AE Enschede, The Netherlands;

³Department of Quality and Process Innovation, Academic Medical Center Amsterdam,
Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands;

⁴ Program leader STREAM (Healthcare Logistics), Roessingh Rehabilitation Center,
Roessinghsbleekweg 33, P.O. Box 310, 7500 AH, Enschede, The Netherlands.
{n.baer,n.kortbeek,n.litvak}@utwente.nl; o.roukens@roessingh.nl

Abstract

Background: Rehabilitation Care is a treatment process that involves multiple disciplines, which are affiliated with different departments and use different planning horizons. This often results in planning difficulties and long waiting lists. In order to increase efficiency the Rehabilitation Centre "Het Roessingh" in the Netherlands has introduced the concept of treatment plans.

Purposes: We support Het Roessingh to make their organizational process ready for implementation of treatment plans with the managerial purpose to prospectively assess the consequences of various interventions, without experimenting on the real-world system.

Methods: We use mathematical models based on techniques from Operations Research and Management Science. A simulation model is used to analyze the integral impact of proposed organizational solutions on the patient flows.

Findings: First, given the number of referrals per week, a necessary number of intake slots is determined in order to satisfy the access time requirements. Second, we quantify capacity losses resulting from cancellations of interdisciplinary meetings,. Third, we conclude that several disciplines are going to form bottlenecks in

patient flow once the treatment plans are implemented. Fourth, a balanced staffing rule is computed for each discipline, by which bottlenecks will be avoided and higher efficiency will be achieved,. Finally, we show that the system cannot function at demand/capacity ratios close to 100% because disciplines affect each other.

Practice Implications: A sufficient number of intake slots must be planned weekly. The interdisciplinary meetings are of great importance and must have top priority for practitioners. When the treatment plans are implemented, the proposed balanced staffing rule enables a higher efficiency than the current staffing rule. It must be accepted by the management that multidisciplinary care cannot function with 100% efficiency, because of the necessity to buffer for variation in demand, and the interdependency between the disciplines in each patient treatment.

Keywords: Health Care Management, Operations Research, Patient Flow Analysis, Markov models, Discrete event simulation.

1 Introduction

Rehabilitation is the process in which a patient is assisted in improving or recovering lost functions after an event, illness or injury that causes functional limitations. The patient is treated by a multidisciplinary team of a rehabilitation physician and therapists for a period of time. Each team member treats the patient in different segments of the rehabilitation process. Therefore, coordination within both the care process and the logistical organization is essential (Kroll & Neri, 2003; Wade & de Jong, 2000). As in many health care processes, planning deficiencies have a negative impact on both quality of care and logistical efficiency (World Health Organization & The World Bank, 2011; Conforti, Guerriero, & Guido, 2008). The multidisciplinary nature makes planning and control of the rehabilitation process a challenging task. In this article, we perform a patient flow analysis of the Dutch rehabilitation centre “Het Roessingh” and address the related resource capacity planning and control issues.

According to the recent WHO report on disability (World Health Organization & The

World Bank, 2011), in high-income countries about 18% of the population lives with some form of disability, while the prevalence of disability is rising due to ageing populations and the global increase in chronic health conditions. The expenditures for rehabilitation care pay off well due to the considerable benefits, consisting of enhanced economical activity, health outcomes, educational achievements, and participation in community activities of people with disabilities. Public spending on disability programmes amounts to 1.2% of GDP for OECD countries, and is particularly high in the Netherlands and Norway, where the expenditures are about 5% of GDP. Both the WHO (World Health Organization & The World Bank, 2011) and the Dutch rehabilitation sector (SEO Economisch Onderzoek & Revalidatie Nederland, 2008) observe a large potential for organizing rehabilitation care more efficiently and effectively. The Dutch rehabilitation centre ‘Het Roessingh’ provides a striking example. While the treatments at its Pain Rehabilitation department are effective (the treatment programs are accredited by the Commission on Accreditation of Rehabilitation Facilities (CARF) International, 2012), its management observes considerable organizational challenges: waiting lists are long (on average patients wait more than 80 days for their first consultation), practitioners experience high working pressure, and the insight into the demand and supply of pain rehabilitation care is insufficient.

To tackle these challenges, inspired by the concept of clinical pathways (de Bleser et al., 2006), Het Roessingh is introducing the concept of ‘treatment plans’. Treatment plans specify the required treatment for specific groups of patients with the same diagnosis during a period of several weeks or months. They intend to ensure that the right treatments are provided at the right time, so that the best quality of care is realized while making efficient use of available resources (Dafoe, Arthur, Stokes, Morrin, & Beaton, 2006).

The methods and analyses presented in this article are based upon Operations Research techniques. They support Het Roessingh to make their organizational process ready for complete implementation of treatment plans with the managerial purpose to prospectively assess the consequences of various alternative interventions, without actu-

ally changing the system. Operations Research techniques have proved to be valuable to assess such effects in health care settings and rationalize decision making (Brandeau, Sainfort, & Pierskalla, 2004; Jun, Jacobson, & Swisher, 1999).

The complexity of the rehabilitation care chain is induced by its multidisciplinary nature and length of the treatments, which results in logistical interactions between the care providers. The developed models will support Het Roessingh to gain insight into the behavior of their care chain, and to obtain insight in demand for care and the capacity required to meet a certain quality of service. The integral patient flow is addressed, from the intake consultation until the end of treatment. Modeling is highly suitable in this problem setting, since experimenting in practice induces risks for patients and field experimenting makes it difficult to control all variables, can raise resistance of stakeholders, takes more time, and offers less statistical reliability. The analysis enables us to advise Het Roessingh on the optimal system configuration of the treatment plan based care chain and to derive rules of thumb that can be applied in its design and control.

Rehabilitation Centre “Het Roessingh”

Het Roessingh employs over 600 persons (400 FTE) performing rehabilitation treatment and medical research. Each year, about 3800 patients are treated in this centre in three different departments: Adult Rehabilitation, Children’s Rehabilitation and Pain Rehabilitation. In this study we focus on the Pain Rehabilitation department. This department treats over 900 patients per year (roughly 53.000 consultations) suffering from chronic pain or chronic fatigue syndrome for which no medical treatment is known. Patients are taught to cope with their pain, thereby trying to enable them to fully participate in society. Clinicians from six medical disciplines are involved in the treatment at Pain Rehabilitation: Rehabilitation Physicians (RP), Physiotherapists (PT), Occupational Therapists (OT), Psychologists (PS), Social Workers (SW), and Kinesiologists (KI).

The current patient flow, before the introduction of treatment plans, is depicted in Figure 1. Patients who are referred to the pain rehabilitation department (17.7 per week

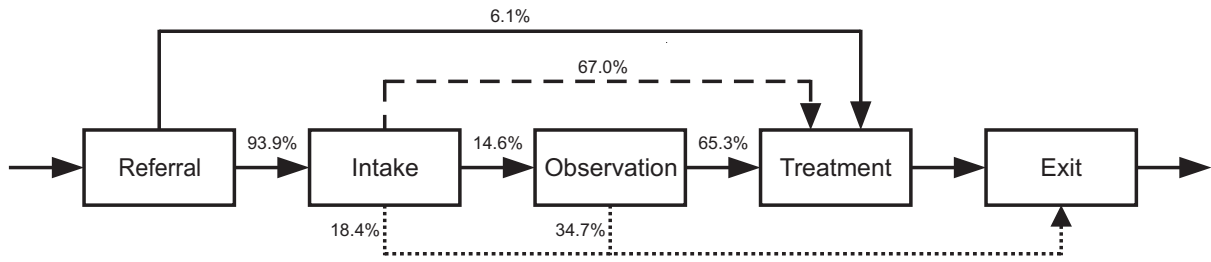


Figure 1: Patient flow diagram in the Pain Rehabilitation department of Het Roessingh.

on average) by a physician, will generally first receive an intake consultation (93.9%). The intake consists of interviews by multiple therapists of different disciplines and a rehabilitation physician. Advised by the multidisciplinary team of therapists, the physician decides whether the patient is eligible for treatment at Het Roessingh (81.6% of the intake patients) and if so which treatment is most suitable. If no proper diagnosis could be made during the intake, the first period of the treatment consists of a so-called “observation period” (14.6% of the intake patients), during which the appropriate follow-up treatment is determined. Two types of patients do not require an intake before their treatment, because their medical condition is already known in detail: patients who receive rehabilitation treatment after a cancer treatment (5%) and patients that require only a specific fysiotherapy treatment (1.1%). Patients can be treated as (semi-)inpatient (62%) or outpatient (38%). (Semi-)inpatients require a bed during (part of) the weeks they are in treatment. The majority of the treatments are organized in groups of 6 to 10 patients. Periodically, the rehabilitation physician and the therapists discuss the patient’s condition during an interdisciplinary meeting, to decide whether the treatment should be continued, adjusted or finished.

Improvement potential can be identified in three main drivers behind the care organization of Het Roessingh: coordination, clinician load and capacity balancing.

Coordination. In current practice, treatment planning is decentralized, which means that all disciplines or even therapists manage their own agendas. This hampers coordination within both the care process and the logistical process. As a consequence, in many cases, a short access time and a simultaneously started treatment cannot be realized.

Also, timely planning of follow-up appointments can be problematic, causing discontinuity of the rehabilitation process. Consequently, certain prescribed treatments may never be realized, since they cannot be scheduled. Besides, without centralized planning, the risk of both undertreatment and overtreatment is observed and management experiences difficulties in effectively controlling the organization.

Clinician load. Several disciplines experience demand overload. As a result, waiting lists are long (months) and growing, work pressure is high, and the total duration of treatments is unpredictable due to the occurring discontinuity. The management indicates that therapists do not have sufficient time for administrative duties related to a patient's treatment, so-called indirect patient care. This has a negative effect on the reporting quality, which results in the cancellation of interdisciplinary meetings where a patient can be discharged, since these can only effectively be performed if the required documentation is available. If an interdisciplinary meeting is cancelled, in general, the treatment is extended at least until the next interdisciplinary meeting. This effect fortifies itself, since unnecessarily extending treatments claims capacity which could otherwise be used to admit new patients from the waiting list.

Capacity planning. Measurements reveal that there exists an imbalance between the capacities that the different disciplines have available. Consequently, while some disciplines experience demand overload, others experience underutilization. This results in a situation where waiting lists are growing, while at the same time available capacity remains unused.

The intervention that Het Roessingh intends to make to address the coordination issue is to introduce treatment plans. Each treatment plan takes a fixed number of weeks. The entire treatment is scheduled in advance, including the interdisciplinary meetings, in which the progress of the patient is discussed. Since a simultaneous start at the different disciplines is essential for a coordinated care process, a patient's treatment is then only allowed to start if each of the required disciplines has capacity available for the complete period of the treatment.

In this article, we study the effect of the introduction of treatment plans on clinician load and capacity balancing issues. In more detail, our contributions are as follows: (i) mapping of the pain rehabilitation process and treatment plans (Section 3.1), (ii) determining the capacity required for intake consultations (Sections 3.2, 4.1), (iii) illustrating the impact of canceling interdisciplinary meetings (Sections 3.3, 4.2), (iv) computing the mean and standard deviation of the demand per discipline (Sections 3.5, 4.3), (v) defining rules of thumb for balancing discipline capacities (Sections 3.6, 4.4), and (vi) evaluating the complete system redesign in terms of quality of service and logistical efficiency (Sections 3.4, 4.4).

2 Conceptual Framework

2.1 OR/MS in health care

The analysis in this article addresses the managerial function of resource capacity planning and control as defined in (Hans, van Houdenhoven, & Hulshof, 2012; Hulshof, Kortbeek, Boucherie, & Hans, 2011): “resource capacity planning and control addresses the dimensioning, planning, scheduling, monitoring, and control of facilities, equipment and staff”, using the methods of Operations Research and Management Science (OR/MS).

OR/MS is an interdisciplinary branch of applied mathematics, engineering and sciences that uses various scientific research-based principles, strategies, and analytical methods including mathematical modeling, statistics and algorithms to improve an organization’s ability to enact rational and meaningful management decisions (INFORMS Website, 2012). Since the 1950s, the application of OR/MS to health care yields significant contributions in accomplishing essential efficiency gains in health care delivery. The topics addressed include operating room planning, nurse staffing and appointment scheduling. Due to the interdisciplinary nature of OR/MS applied to health care, there is an extensive base of literature published across various academic fields. We refer to the online literature database ‘ORchestra’ (Hulshof et al., 2011), in which references in

the field of OR/MS in health care are categorized by medical and mathematical subject.

2.2 Treatment plans

Het Roessingh intends to treat all patients according to so-called treatment plans. Treatment plans resemble the concept of clinical pathways or critical pathways, but are in fact not the same. Although a unique definition of clinical pathways is lacking (Vanhaecht, de Witte, Depreitere, & Sermeus, 2006), the definition given by Medical Subject Headings (MESH) of Pubmed (Medical Subject Headings (MeSH), 2012) is: “schedules of medical and nursing procedures, including diagnostic tests, medications, and consultations designed to effect an efficient, coordinated program of treatment.” A treatment plan differs from a clinical pathway as it is used as a blueprint for a patient’s treatment for some weeks or even months, where a clinical pathway typically prescribes in detail all required activities during the time horizon of a number of days. The treatment plans at Het Roessingh have been created in a similar way as clinical pathways. A thorough description of clinical pathways, their creation and their usage can be found in (Campbell, Hotchkiss, Bradshaw, & Porteous, 1998). A literature review on the usage of clinical pathways in clarifying patient flows is given in (Vanberkel, Boucherie, Hans, Hurink, & Litvak, 2010).

The introduction of treatment plans has both a medical and a logistical motivation, as it creates clarity for both patients and practitioners. It realizes uniformity in the care process, so that the risk of both undertreatment and overtreatment are minimized (SEO Economisch Onderzoek & Revalidatie Nederland, 2008). Also, it prevents discontinuity in the care process and it stimulates the coordination among the different involved practitioners (Braaksma, Kortbeek, Post, & Nollet, 2012). From a planning perspective, the treatment plan concept offers insight in required capacity for the patient mix the facility serves. In addition, instead of scheduling a patient treatment one week at a time, a treatment plan offers the opportunity to schedule the total treatment at once.

2.3 Staffing and scheduling for treatment plans

Appointment scheduling in health care is a topic that is well-addressed in OR/MS literature, see e.g. (Cayirli & Veral, 2003; Gupta & Denton, 2008) for comprehensive surveys. However, the literature has mostly focused on scheduling a given number of single appointments on a particular day for an individual service provider (Cayirli & Veral, 2003; Chien, Huang, & Hu, 2009; Chien, Tseng, & Chen, 2008). Scheduling multiple appointments for monodisciplinary treatment for a time horizon of a week or several weeks is addressed in (Ogulata, Koyuncu, & Karakas, 2008) for physiotherapy, in (Conforti et al., 2008; Conforti, Guerriero, Guido, & Veltri, 2011) for radiotherapy, and in (Turkcan, Zeng, & Lawley, 2010) for chemotherapy. In Braaksma et al., 2012, an algorithm is presented to schedule an entire multidisciplinary rehabilitation treatment at the moment a patient treatment request is placed. This reference does not focus on staffing: the discipline capacities are taken as input for the scheduling algorithm. Staffing is an important aspect in health care, see the literature reviews (Buchan & Dal Poz, 2002; Currie, Harvey, West, McKenna, & Keeney, 2005; Lang, Hodge, Olson, Romano, & Kravitz, 2004). Reviews of OR/MS literature on staffing for health care can be found in (Burke, De Causmaecker, Berghe, & Van Landeghem, 2004; Ernst, Jiang, Krishnamoorthy, & Sier, 2004). To the best of our knowledge, OR/MS references on multidisciplinary staffing for treatment plans are not available.

3 Methods

There are in total 18 treatment plans which cover 6 disciplines and may last up to 39 weeks, which makes the system under study complex. For data collection and process mapping we have used the data systems in Het Roessingh and interviews with the staff including doctors, managers and IT specialists. In the analysis we have used mathematical techniques that in the last years have been successfully employed for improvement of healthcare operations under uncertainty: queueing theory, Markov models, and discrete

event simulations.

An important attribute of our approach is *decomposition* of the process into simpler processes where possible. In particular, we develop three mathematical models that address two specific aspects of the patient flows that can be studied independently. 1) Intakes happen at the beginning of the treatment plan and do not depend on the future treatment. Therefore, we analyze the required number of intakes per week with a separate queueing model. 2) Cancellation of interdisciplinary meetings, where patients are discussed and discharged, result in sometimes unnecessarily prolonged treatment plans. We investigate the effect of such prolongation for each treatment plan separately. 3) We develop a Markov model that describes the progress of patients through treatment plans. The proposed mathematical models are then combined in a simulation model. In the following subsections we describe the methods in more detail.

3.1 Data analysis and process mapping

First, we estimate a patient's path through (usually multiple) treatment plans. The treatment plans created by Het Roessingh are blueprints. Some treatments in treatment plans are given only to a fraction of patients, and the duration of the appointments can vary from one patient to another.

We have used historic data and expert opinion to obtain the required information. This resulted in 18 treatment plan blueprints (examples are oncology, whiplash, and back pain). Each treatment plan blueprint prescribes the disciplines a patient of a particular type should be treated by, the required number of treatments per discipline, the duration of each treatment (in minutes) and the week number in which it should take place. In addition, the fractions are included according to which particular appointments series are required for a specific patient, so that individualized treatment plans can be realized based on the treatment plan blueprints. The results for one of the two observation treatment plans are presented in Table 1. This treatment plan of five weeks involves six disciplines.

Next, we analysed *sequential treatment plans and transition percentages*. Usually a

Discipline	Appointment	Group size	Treatment duration					Required for $x\%$
			Week					
			1	2	3	4	5	
Rehabilitation physician (RP)	Therapy 1	1	20	20	20	-	-	100
	Therapy 2	1	-	-	-	-	30	100
	Interdisciplinary Meeting	1	15	15	15	15	-	100
Physiotherapist (PT)	Therapy 1	1	60	60	60	-	-	100
	Interdisciplinary Meeting	1	15	15	15	15	-	100
Occupational therapist (OT)	Therapy 1	1	60	60	60	-	-	75
	Interdisciplinary Meeting	1	15	15	15	15	-	100
Psychologist (PS)	Therapy 1	1	60	60	60	-	-	100
	Therapy 2	1	60	60	-	-	-	80
	Therapy 3	1	-	-	-	-	30	100
	Interdisciplinary Meeting	1	15	15	15	15	-	100
Social worker (SW)	Therapy 1	1	60	60	60	-	-	80
	Therapy 2	1	60	-	-	-	-	100
	Interdisciplinary Meeting	1	15	15	15	15	-	100
Kinesiologist (KI)	Therapy 1	3	30	30	30	-	-	100
	Therapy 2	3	30	30	30	-	-	100
	Therapy 3	1	30	30	30	-	-	75
	Interdisciplinary Meeting	1	15	15	15	15	-	100

Table 1: Example of a treatment plan blueprint (with the treatment durations given in minutes).

treatment plan is only a part of a complete care process. For example, "Intake" and "Observation" are treatment plans, which are usually followed by other treatment plans. Thus, we need to determine transition percentages between treatment plans. For instance, if the transition percentage from treatment plan 1 to treatment plan 2 is 15%, then 15% of all patients who completed treatment plan 1 will continue with treatment plan 2. Since treatment plans are not yet applied in practice, the transition percentages have been estimated through historic data on the current process. The transition percentages were recorded in 18×18 matrix Q , where entry (j, k) , $Q_{(j,k)}$, is a percentage that treatment plan j is followed by treatment plan k .

Finally, we establish *priority rules*. By arrival, after intake, and between treatment plans patients can be put on a waiting list if capacity is insufficient to treat them immediately. Priority rules determine which patients will be treated first if capacity is insufficient to treat all patients on the waiting list. Het Roessingh uses the following order of priority: 1) intake; 2) observation; 3) group treatments; 4) individual treatments.

3.2 Number of intakes

Intakes take place for several patients simultaneously because intakes usually require similar sets of specialists and procedures. That is, each week K intake sessions are held, each consisting of N patients. Intake sessions take place only when they are completely filled; if the number of patients waiting for intake is less than N the intake session is cancelled. Denote the average number of new patients registered in one week by λ . Then the questions that are of interest for Het Roessingh can be formulated as follows. Given λ , K and N , what is the average access time for an intake and what is the probability that the access time exceeds a specified norm? We answer these questions using a queueing model, which is presented in Appendix A.

3.3 Interdisciplinary meetings

At fixed points in time, at the interdisciplinary meetings for individual outpatients, specialists discuss the progress of a patient and continuation of the treatment. The first meeting takes place after three weeks, and subsequently every sixth week. When such a meeting is cancelled then usually the treatment period of a patient is extended. Practice shows that the last meeting in the treatment plan is crucial, and whenever this meeting is cancelled the treatment will definitely be extended. For any other meeting the choice of extending the treatment depends highly on the patient's condition. When a treatment is extended, it will be extended with one period of six weeks and one interdisciplinary meeting.

In practice, there is no formal policy for treatment extensions. We consider two possible situations: 1) each cancelled meeting leads to an extension of the treatment; 2) only cancellation of the last meeting leads to an extended treatment. After introduction of the treatment plans, the practical situation will be in between the two options. Applying a simple mathematical method (presented in Appendix B) reveals the negative effect of cancelled meetings on treatment durations, see Section 4.2.

3.4 Patient flows

We design a so-called Markov model that describes possible treatment scenarios of a patient (see Appendix C). The required input for the model is estimated from historical data. We use this model to obtain the capacity requirements, that is, the total treatment time (in Full Time Equivalents) for each discipline per week. These are further used to find bottlenecks in the patient flows and can also be used to determine a staffing rule.

3.5 Capacity requirements

We determine the average capacity requirements of the network of patient flows induced by the treatment plans, by (1) calculating the average capacity in hours required per appointment of each type in each week, (2) computing the average capacity required per patient in each week of treatment plan j , and (3) considering all possible sequences of treatment plans $j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_k \rightarrow j$ that lead to the treatment plan j . The network of treatment plans is such that no patient receives a particular treatment plan twice. Combining (1), (2), and (3) with the average number of arrivals per week gives us for each discipline the average capacity required to treat all patients. The mathematical formulation of this method is presented in Appendix D.

3.6 Staffing improvements

In order to balance the load of disciplines and to improve the performance in terms of number of treated patients, we propose to determine the staffing level such that each discipline has the same (theoretical) load ρ^d , defined as the average capacity requirement for the discipline divided by the available capacity for this discipline (see Appendix E for details). It is well-known from queueing theory that the load of the system is a defining parameter for the magnitude of queue lengths, access times, and probability of no queue. Therefore, by leveling the load between disciplines, we achieve comparable waiting times for different classes of patients.

3.7 Simulation model

We investigate the performance under different load and staffing levels using discrete event simulation of patient flows (Law, Kelton, et al., 2005). The **input** for the simulation model consists of: the treatment plans, the transition matrix Q , the priority rules, the available capacity (either the current staffing rule or our proposed staffing rule), and the mean number of weekly referrals. The **process** consists of the discrete event simulations that mimic, per week, the entire network of patient flows within the Pain Rehabilitation department in Het Roessingh. The number of arrivals in a week is modelled as a Poisson random variable. The simulation model then performs two actions: 1) for each new patient a first treatment plan is determined using the transition matrix Q ; 2) each new patient is put on the waiting list corresponding to his/her treatment plan. For each week the simulation model checks whether patients on the waiting lists can start their treatment or not. If there is enough capacity available to treat a patient then the patient starts the treatment, and the available amount of capacity is decreased by the corresponding amount of hours, specified by the treatment plan. Priority rules (see Section 3.1) are used to determine the order in which the patients from the waiting list receive their treatment. When a patient has completed his treatment plan then there are two options: either the patient starts a new treatment plan (which is determined using the transition matrix Q) or the patient leaves the system. The **output** consists of several performance measures: the time a patient must wait before an intake is completed or until the treatment has started, the total duration of a patient's treatment, and the load or utilization of each specific discipline.

4 Findings

Our models enable Het Roessingh to evaluate various improvements of both quality of care and efficiency, ranging from organizing intake procedures and prioritizing interdisciplinary meetings, to indicating bottlenecks and determining staffing levels. This section presents some highlights.

4.1 Number of intakes

Dutch health care organizations have agreed upon an access time norm of seeing 80% of the patients within three weeks (Treek Norm, 2012). In addition, it is a target to see all patients within four weeks. In the current configuration Het Roessingh has $K = 3$ intake sessions of group size $N = 5$, which is insufficient to handle the average number of referrals per week ($\lambda = 17.7$). Our results show that increasing the number of intakes to $K = 3$ and $N = 6$, or $K = 4$ and $N = 5$, will be sufficient to satisfy this norm.

Figure 2a gives the average access time for increasing λ , resulting from the queueing model of Section 3.2. Observe that two situations lead to extremely long waiting lists and access times. Firstly, when $\lambda \approx K \cdot N$, the load of the system approaches 100%, which results in extremely long access times. Secondly, when $\lambda \approx 0$ it is difficult to form a group of N patients to plan an intake session because there are almost no arrivals, and thus access times are long. Clearly, the first situation is especially relevant to Het Roessingh. Figure 2b displays the access time percentiles for configurations $K = 3, N = 6$ and $K = 4, N = 5$. The configuration $K = 3$ and $N = 6$ is not indicated in Figure 2b as it gives 0% seen within four weeks.

For $\lambda = 17.7$, we advise to use $K = 4$ and $N = 5$, and the access time norm is satisfied, all patients are seen within four weeks ($\gg 99.9\%$), and the utilization of the intake sessions is high: 88.5%.

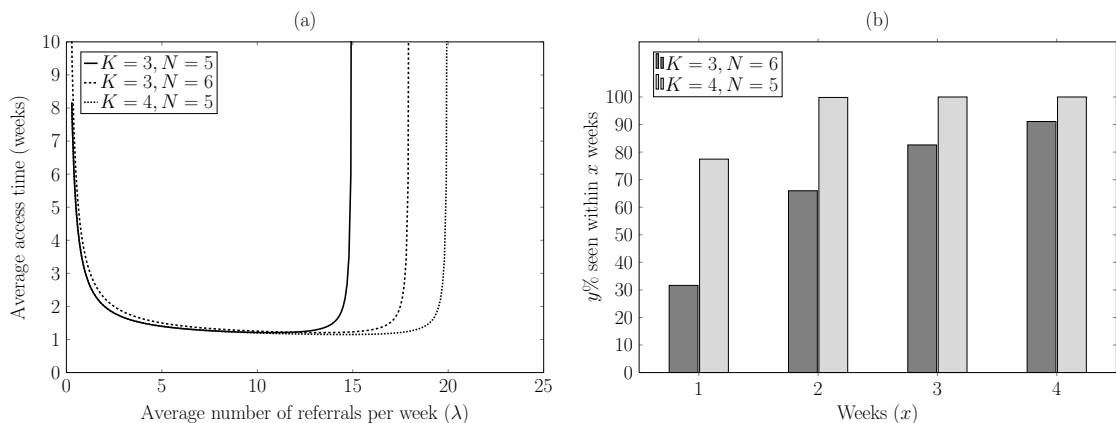


Figure 2: Intake access time: (a) average, depending on λ , and (b) percentiles per week for $\lambda = 17.7$.

4.2 Interdisciplinary meetings

Consider a treatment plan of 15 weeks that involves three interdisciplinary meetings. Let p denote the fraction of cancelled meetings. The average extensions of the treatment duration under the two scenarios of Section 3.3 are as follows:

- *Scenario 1* (an additional six treatment weeks for every cancelled interdisciplinary meeting): $\frac{18p}{1-p}$.
- *Scenario 2* (an additional six weeks if the last interdisciplinary meeting is cancelled): $\frac{6p}{1-p}$.

The cancellation fraction p has a significant effect on the length of the treatment plan, as illustrated in Figure 3. For instance, under scenario 2, for $p = 0.1$ the average extension is 0.67 weeks, while for $p = 0.4$ it increases to 4 weeks. Although the clinicians usually give a higher priority to patient treatments than to the interdisciplinary meetings, our results show that the management must facilitate a higher attendance of the meetings, otherwise the scarce capacity will be spend, in large amounts, on patients that could have been discharged, while the patients who need care will be placed on the waiting list.

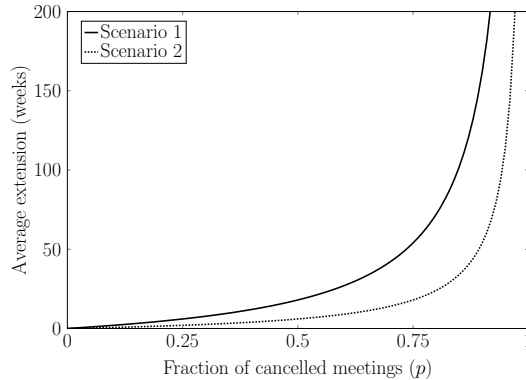


Figure 3: An illustration of the effect of cancelling interdisciplinary meeting.

4.3 Capacity Requirements and bottlenecks

Evaluating the ability of Het Roessingh to accommodate the implementation of treatment plans with the current discipline capacities, requires the identification of bottlenecks

and of the maximum number of patients that can be treated by each discipline. The Markov model from Sections 3.4 and 3.5 allows us to identify the bottleneck disciplines, and to determine the maximum number of patient referrals each discipline can handle, by calculating the average capacity requirements per patient referral. The results are presented in Table 2. The capacity requirements are measured in Full Time Equivalents (FTE's).

We indicate the bottlenecks in patient flows by comparing the capacity requirements to the actual staffing, for each discipline. When the average capacity requirement exceeds the available capacity, it is a clear indication that the discipline d constitutes a bottleneck in the system. The fifth column shows that almost all disciplines are overloaded, since the ratio demand over capacity exceeds one. Also, we see that the discipline Occupational Therapists can only handle an average of 8 referrals per week. If we compare this with the actual number of referrals per week (which is on average 17.7) it is clear that the discipline OT is the main bottleneck for the network of patient flows induced by the treatment plans. The conclusion is that complete implementation of treatment plans is only possible if a lower number of referrals is allowed or if staffing levels are adjusted.

Discipline	Current Staffing (S)	Average capacity required per referral (A)	Capacity required for $\lambda = 17.7$ (D)	Load under $\lambda = 17.7$ (=D/S)	Maximum throughput (=S/A)
RP	4.00	0.142	2.509	0.63	28.72
PT	7.38	0.859	15.212	2.06	8.59
OT	3.55	0.444	7.850	2.21	8.00
PS	4.53	0.398	7.048	1.56	11.38
SW	4.69	0.369	6.531	1.39	12.71
KI	2.58	0.151	2.664	1.03	17.14
Total	26.73	2.363	41.820	1.56	8.00

Table 2: Capacity requirements (in FTE) and bottleneck identification.

4.4 Staffing Improvements

Het Roessingh strives to uniformize its processes, switch to treatment plans, reduce the work pressure, and avoid ad hoc decisions in the future. Our results help to achieve these goals by having revealed which disciplines in Pain Rehabilitation suffer from the highest overload in the desirable system design, and by proposing staffing improvements.

To illustrate imbalance of the current staffing and to identify the potential of improved staffing rules as introduced in Section 3.6, we use the simulation model from Section 3.7 to analyse system performance under various referral rates and staffing rules. For each scenario, we perform 15 simulation runs, each simulating 3000 weeks in the Pain Rehabilitation department, of which 500 weeks are taken as warm-up period. Applying Welch's procedure (Welch, 1981), these settings provide a half-width of 5% for the 95% confidence intervals for the various performance indicators. The results are shown in Table 3.

We predict the performance of the system under its new design, which is different from current practice in several fundamental ways. First, the treatment plans have not yet been implemented in Het Roessingh. Second, no ad hoc flexibility in prioritizing patients will be allowed other than specified by type (i.e. intake, observation, group, individuals). Third, only allows to schedule a patient when capacity is available with all disciplines for his/her entire treatment plan. Currently, only part of the treatment plan is scheduled at a time, and ad hoc adjustments are applied for patients who have been waiting too long, so that they receive a higher priority. Finally, indirect times will be included (reporting, administration) in the treatment time while in current practice it is possible to fill out doctors' schedules with treatments only.

The third column of Table 3 represents the current staffing. In Section 4.3, we have shown that multiple disciplines are overloaded when $\lambda = 17.7$, thus the queueing system under study is overloaded, implying that queue lengths and access times will grow rapidly without boundaries. Table 2 shows that under the current staffing, an upper bound on the maximum referral rate is 8.0 per week, dictated by the OT discipline. The inherent variability of demand and the interdependence between disciplines makes that that the referral rate must be even lower to keep access times and treatment delays stable. Experimenting with different referral rates revealed that $\lambda = 7.5$ would be manageable. With this load, we see in the second column of Table 3 that the utilization per discipline is varying between 26% and 93%, which is an indication of unbalanced staffing. We note that in this case the waiting times for individual patients are highly fluctuating. Since

	$\lambda \rightarrow$	7.5	7.5	7.5	17.7	17.7	17.7
	Staffing Rule \rightarrow	current	80%	85%	80%	85%	90%
Staffing (in FTE)							
RP		4.00	1.33	1.25	3.14	2.95	2.79
PT		7.38	8.06	7.58	19.02	17.90	16.90
OT		3.55	4.16	3.92	9.82	9.24	8.73
PS		4.53	3.73	3.51	8.81	8.29	7.83
SW		4.69	3.46	3.26	8.16	7.68	7.26
KI		2.58	1.41	1.33	3.33	3.13	2.96
Total		26.73	22.15	20.85	52.28	49.20	46.47
Load							
RP		26.5%	79.9%	85.2%	80.0%	85.0%	90.0%
PT		87.2%	79.9%	85.1%	79.8%	85.1%	90.1%
OT		93.7%	79.8%	85.1%	79.9%	85.0%	90.0%
PS		65.8%	79.9%	85.2%	79.9%	85.0%	89.9%
SW		58.8%	79.8%	85.1%	79.9%	85.1%	90.0%
KI		43.6%	79.8%	85.2%	80.0%	85.1%	89.9%
Total		66.2%	79.9%	85.1%	79.9%	85.1%	90.0%
Average (in weeks)							
Access Time Intake		1.4	1.4	1.5	1.1	1.1	1.2
Access Time Individual		2.2	0.2	0.7	0.0	0.1	0.7
Access Time Group		9.9	10.4	13.2	4.0	4.3	8.3
Access Time Observation		1.8	1.9	2.5	0.7	0.8	1.3
Prescribed Treatment duration		21.3	21.3	21.3	21.3	21.3	21.2
Delay within Treatment		4.1	3.5	4.8	1.3	1.4	3.3
90-th Percentile (in weeks)							
Access Time Intake		2.0	2.0	2.0	2.0	2.0	2.0
Access Time Individual		5.9	1.0	2.2	0.0	0.8	2.2
Access Time Group		22.6	23.8	28.1	9.4	10.0	14.9
Access Time Observation		5.0	5.0	6.3	2.0	2.0	3.3
Prescribed Treatment duration		36.0	36.0	36.0	36.0	36.0	36.0
Delay within Treatment		11.8	10.5	13.5	4.0	4.0	6.9

Table 3: The results for multiple staffing rules (intake setting: $K = 4, N = 5$).

scheduling individual treatments receives the lowest priority, this can be regarded as a sign that the system is operating close to its maximum capacity.

In columns 4 and 5 the results from aligning the number of FTEs per discipline to demand according to our proposed staffing rule with $\lambda = 7.5$ are shown, with the preset load 80% and 85%. We see that the average access times and total treatment times are comparable with the third column but that the load per discipline is better distributed and higher on average. Our proposed staffing rule reduces the total number of FTE by 17% to 22% without decreasing system performance. Under a preset load per discipline of 90%, the system is overloaded. This is explained by the fact that disciplines are highly dependent: a discipline can often not start a treatment for one patient before the treatment of another patient is completed by another discipline. Therefore, some capacity loss is unavoidable. Here we also clearly observe the economies of scales effect: when $\lambda = 17.7$, the system can function under the preset load 90% per discipline.

In the last three columns the results are presented using our staffing rule for $\lambda = 17.7$, with preset loads of 80%, 85% and 90%. We see that the loads per discipline are balanced and access times are small. Treatment delays decrease compared to the case with $\lambda = 7.5$, so that treatment discontinuity is reduced. A final apparent observation is that to be able to fully implement treatment plans under a demand scenario of 17.7 referrals per week, a significant increase in the number of FTEs is required.

5 Practice Implications

The results of this case study can be subdivided in two categories: basic insights and tools. The basic insights are the evaluation of the required number of intakes per week, and the quantitative characterization of the importance of interdisciplinary meetings. In Section 4.1, we have shown that there is flexibility in how intake sessions can be organized, but the average number of available intake slots per week must be larger than an average number of weekly referrals by a visible margin, otherwise, the access times will inevitably increase beyond the Dutch national access time norm. To this end, we advise Het Roessingh to plan four intake sessions during each of which five patients can be seen. One could decide to plan more intake sessions but this will result in a decrease of the utilization of the planned intake sessions. In Section 4.2, we have demonstrated that cancellations of interdisciplinary meetings lead to unnecessary capacity losses. We conclude that the interdisciplinary meetings should have top priority, and the management must facilitate a high attendance of these meetings by the doctors and therapists.

We have developed tools to determine capacity requirements (Section 4.3) and to formulate staffing rules (Section 4.4). These tools are tailored to the intended implementation of treatment plans. Indirect time is included in the treatment time to avoid clinician overload, and to ensure that interdisciplinary meetings are not cancelled as a result of unavailability of the required documentation. Our methods allow to evaluate

system design before actual changes are made. Therefore, they can assist the management in making optimal logistical decisions during the implementation of treatment plans.

To make patient flows more predictable, and to offer patients a more consistent and reliable quality of care, more and more rehabilitation centers organize treatments according to treatment plans. We have shown that OR/MS methods, being able to quantitatively evaluate patients flows and prospectively assess the system's performance, are most suitable to support this organizational change. One of the main challenges of this study has been to map the treatment plans from the available data. In this article, we have described what type of data is essential for mapping, modelling and evaluating the patient flows. Our proposed methodology for the treatment plans mapping and system analysis, developed for Het Roessigh, is general which makes that it can be extended and applied in other rehabilitation centers.

Appendix

A Model for the intake process

This appendix belongs to Section 3.2. Let L_t denote the number of patients waiting for an intake at the beginning of week t , $t = 1, 2, \dots$. Next, let A_t be the number of new patients (arrivals) in week t , and S_t the number of intakes (services) in week t . Then the queue length L_t satisfies a so-called Lindley's recursion:

$$L_{t+1} = L_t + A_t - S_t, \quad t = 1, 2, \dots \quad (1)$$

Indeed, compared to the queue length at the beginning of week t , the number of arrivals must be added and the number of completed intakes must be subtracted in order to obtain the queue length at the beginning of week $t + 1$. Furthermore, we note that S_t depends on L_t as follows: $S_t = \min \{N \cdot \lfloor \frac{L_t}{N} \rfloor, K \cdot N\}, t = 1, 2, \dots$. This expression reflects the maximum number of sessions that can be filled, which has an upper limit of

K , multiplied by the size of each session N . Assume that the number of arrivals in each week is independent of each other. Then $\{L_t, t \geq 1\}$ is a Markov process, of which the stationary probabilities, $\pi_l = \lim_{t \rightarrow \infty} P(L_t = l)$, $l \geq 0$, can be found.

We model A_t , $t = 1, 2, \dots$, as independent random variables, each having a Poisson distribution with mean λ . The assumption of Poisson distribution is common and most suitable if patients arrive independent of each other, as is the case in our study. We choose a truncation approximation, to approximate equation (1). Specifically, we assume that L_t cannot exceed some large but finite number l_{max} and numerically solve the resulting finite Markov chain. The obtained stationary probabilities provide a good approximation for the π_l 's, and are then used to find the stationary probability distribution of the access time.

B Model for interdisciplinary meeting cancellations

This appendix belongs to Section 3.3. Let p denote the probability that an arbitrary interdisciplinary meeting is cancelled. For convenience of the mathematical presentation, we first present situation 2, and then situation 1.

Situation 2. If only the last meeting is important then extension of at least one period happens with probability p , and the probability of k six-week-extensions is $p^k(1 - p)$, $k = 0, 1, \dots$. This is a shifted geometric distribution with mean $p/(1 - p)$.

Situation 1. If each cancelled meeting leads to a six-weeks prolongation, then the treatment continues until all planned meetings have taken place. Since the treatment plan contains $x \geq 1$ meetings, the probability distribution of the number of six-weeks-extensions is a sum of x random variables with a shifted geometric distribution with mean $p/(1 - p)$.

C Model for patient flows

This appendix belongs to Section 3.4. Let l_j be the duration in weeks of treatment plan $j = 1, 2, \dots, 18$. For each week $w = 1, 2, \dots, l_j$ in a treatment plan j , for each discipline $d = 1, 2, \dots, 10$, we are given a complete description of the number of appointments of type a , $n_{w,j}^{d,a}$, and the probability that an appointment is needed, $p_{w,j}^{d,a}$. Our modelling assumption is that transitions between treatment plans happen according to a Markov process, that is, after completing treatment plan j a patient proceeds to treatment plan k with probability $Q_{(j,k)}$, independent of the past.

D Model for capacity requirements

This appendix belongs to Section 3.4. The calculations presented here, are performed for each discipline separately, therefore we omit a superscript d . Let $L_{w,j}^a$ denote the random appointment length, and $Z_{w,j}^a$ the mean amount of capacity (in hours) needed per appointment of type a in week w . Then, clearly, $\mathbb{E}[Z_{w,j}^a] = n_{w,j}^a \cdot \mathbb{E}[L_{w,j}^a] \cdot p_{w,j}^a$, and the mean amount of capacity needed per patient in week w of treatment plan j is $\mathbb{E}[Z_{w,j}] = \sum_a \mathbb{E}[Z_{w,j}^a]$.

Denote by $N_{w,j}$ the number of patients that, at a certain point in time, are in week w of a treatment plan j . Using the transition matrix and the mean number of arrivals per week, we are able to determine $\mathbb{E}[N_{w,j}]$ as follows. Consider all possible sequences of treatment plans $j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_k \rightarrow j$ that lead to the treatment plan j . Assume also that all interdisciplinary meetings take place as planned, so there is no random prolongation of the treatment plans. Recall that λ is the average number of newly arrived patients per week. Then for each $w = 1, \dots, l_j$, $j = 1, 2, \dots, 18$, we have

$$\mathbb{E}[N_{w,j}] = \sum_{0 \rightarrow j_1 \rightarrow j_2 \rightarrow \dots \rightarrow j_k \rightarrow j} \lambda \cdot Q_{(0,j_1)} Q_{(j_1,j_2)} Q_{(j_2,j_3)} \cdots Q_{(j_{k-1},j_k)} Q_{(j_k,j)}.$$

Denote by $C_{w,j}$ the amount of capacity needed in week w for treatment plan j . Under the

natural assumption that the numbers $N_{w,j}$ and $Z_{w,j}$ are independent, we find $\mathbb{E}[C_{w,j}] = \mathbb{E}[Z_{w,j}]\mathbb{E}[N_{w,j}]$. The mean amount of capacity needed to treat all patients, the *capacity requirements* for a discipline is given by $\mathbb{E}[C] = \sum_{j=1}^{18} \sum_{w=1}^{l_j} \mathbb{E}[C_{w,j}]$. This approach can be extended to the case when interdisciplinary meetings are sometimes cancelled leading to longer treatments, by allowing a random length of a treatment plan.

E Model for staffing improvements

This appendix belongs to Section 3.6. To achieve balanced staffing, we set the capacities of the disciplines such that the following holds for each discipline d :

$$\rho^d = \mathbb{E}[C^d]/[\text{available capacity of discipline } d] = \rho.$$

In the context of call-centers and other applications a so-called square-root staffing rule is often used to define the optimal capacity level S : $S = \lambda + \beta\sigma$, where λ and σ are, respectively, the mean and the standard deviation of the workload offered to the system per time unit. This so-called square-root staffing rule is theoretically justified when the amount of work λ grows large, and σ is of the order $\sqrt{\lambda}$, hence, the utilization or load ρ approaches 1 (heavy-traffic regime), see e.g. (Borst, Mandelbaum, & Reiman, 2004). In our situation, however, simulations show that the utilization of almost 100% cannot be achieved due to the interaction between the disciplines (see Section 4.3). Then the square-root staffing rule does not give advantages against the staffing rule with equal ρ that we have chosen. This is confirmed by our numerical results where we computed $\sigma[C^d]$ using similar but more involved derivations as in Appendix D, and then evaluated staffing levels with different values of β . Since we did not find improvements against the staffing rule with equal ρ , we choose to not present the results in this paper.

Acknowledgement

This research is supported by the Dutch Technology Foundation STW, applied science division of NWO and the Technology Program of the Ministry of Economic Affairs, and by the the Royal Dutch Academy for Sciences KNAW.

References

INFORMS Website. (2012).

Retrieved June 21, 2012, from: <http://www.informs.org/>.

SEO Economisch Onderzoek & Revalidatie Nederland. (2008). *Kosten en baten van revalidatie*. (Available at: <http://www.revalidatienederland.nl/actueel/publicaties>)

World Health Organization & The World Bank. (2011). *World report on disability*.

Borst, S., Mandelbaum, A., & Reiman, M. (2004). Dimensioning large call centers. *Operations Research*, 52(1), 17-34.

Braaksma, A., Kortbeek, N., Post, G., & Nollet, F. (2012). Integral multidisciplinary rehabilitation planning. *Technical report, University of Twente, submitted to HCMS, available on request*.

Brandeau, M., Sainfort, F., & Pierskalla, W. (Eds.). (2004). *Operations Research and Health Care: a Handbook of Methods and Applications*. Kluwer Academic Publishers.

Buchan, J., & Dal Poz, M. (2002). Skill mix in the health care workforce: Reviewing the evidence. *Bull World Health Organ*, 80(7), 575-580.

Burke, E., De Causmaecker, P., Berghe, G., & Van Landeghem, H. (2004). The state of the art of nurse rostering. *journal of scheduling*, 7(6), 441-499.

Campbell, H., Hotchkiss, R., Bradshaw, N., & Porteous, M. (1998). Integrated care pathways. *British Medical journal*, 316, 133-137.

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

- Chien, C., Huang, Y., & Hu, C. (2009). A hybrid approach of data mining and genetic algorithms for rehabilitation scheduling. *International journal of Manufacturing Technology and Management*, 16(1), 76–100.
- Chien, C., Tseng, F., & Chen, C. (2008). An evolutionary approach to rehabilitation patient scheduling: A case study. *European journal of Operational Research*, 189(3), 1234-1253.
- Commission on Accreditation of Rehabilitation Facilities (CARF) International. (2012). Retrieved June 21, 2012, from: <http://www.carf.org/home/>.
- Conforti, D., Guerriero, F., & Guido, R. (2008). Optimization models for radiotherapy patient scheduling. *4OR: A Quarterly journal of Operations Research*, 6(3), 263-278.
- Conforti, D., Guerriero, F., Guido, R., & Veltri, M. (2011). An optimal decision-making approach for the management of radiotherapy patients. *OR Spectrum*, 33(1), 123–148.
- Currie, V., Harvey, G., West, E., McKenna, H., & Keeney, S. (2005). Relationship between quality of care, staffing levels, skill mix and nurse autonomy: Literature review. *Journal of Advanced Nursing*, 51, 73-82.
- Dafoe, W., Arthur, H., Stokes, H., Morrin, L., & Beaton, L. (2006). Canadian cardiovascular society access to care working group on cardiac rehabilitation. Universal access: But when? Treating the right patient at the right time: Access to cardiac rehabilitation. *The Canadian journal of Cardiology*, 22(11), 905–911.
- de Bleser, L., Depreitere, R., de Waele, K., Vanhaecht, K., Vlayen, J., & Sermeus, W. (2006). Defining pathways. *Journal of Nursing Management*, 14(7), 553–563.
- Ernst, A., Jiang, H., Krishnamoorthy, M., & Sier, D. (2004). Staff scheduling and rostering: A review of applications, methods and models. *European journal of operational research*, 153(1), 3–27.
- Gupta, D., & Denton, B. (2008). Appointment scheduling in health care: Challenges and opportunities. *IIE Transactions*, 40(9), 800–819.

- Hans, E. W., van Houdenhoven, M., & Hulshof, P. J. (2012). A framework for healthcare planning and control. In R. Hall (Ed.), *Handbook of healthcare system scheduling* (Vol. 168, pp. 303–320). Berlin, Germany: Springer.
- Hulshof, P. J. H., Boucherie, R. J., van Essen, J. T., Hans, E. W., Hurink, J. L., Kortbeek, N., ... Zonderland, M. E. (2011). ORchestra: an online reference database of OR/MS literature in health care. *Health Care Management Science*, *14*(4), 383–384.
- Hulshof, P. J. H., Kortbeek, N., Boucherie, R. J., & Hans, E. W. (2011). *Taxonomic classification of planning decisions in health care: a review of the state of the art in OR/MS* (No. 1944). Enschede: Department of Applied Mathematics, University of Twente. Retrieved from <http://doc.utwente.nl/77509/>
- Jun, J., Jacobson, S., & Swisher, J. (1999). Application of discrete-event simulation in health care clinics: A survey. *Journal of the Operational Research Society*, *50*(2), 109–123.
- Kroll, T., & Neri, M. (2003). Experiences with care co-ordination among people with cerebral palsy, multiple sclerosis, or spinal cord injury. *Disability & Rehabilitation*, *25*(19), 1106–1114.
- Lang, T., Hodge, M., Olson, V., Romano, P., & Kravitz, R. (2004). Nurse-patient ratios: A systematic review on the effects of nurse staffing on patient, nurse employee, and hospital outcomes. *Journal of Nursing Administration*, *34*(7-8), 326-337.
- Law, A., Kelton, W., et al. (2005). *Simulation modeling and analysis* (Vol. 3). McGraw-Hill.
- Medical Subject Headings (MeSH). (2012).
Retrieved June 21, 2012, from: <http://www.nlm.nih.gov/mesh/>.
- Ogulata, S., Koyuncu, M., & 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.
- Treek Norm. (2012).

Retrieved June 21, 2012, from: <http://http://www.treeknorm.nl/>.

- Turkcan, A., Zeng, B., & Lawley, M. (2010). Chemotherapy operations planning and scheduling. *Optimization Online*. Feb, 12.
- Vanberkel, P. T., Boucherie, R. J., Hans, E. W., Hurink, J. L., & Litvak, N. (2010). A survey of health care models that encompass multiple departments. *International journal of Health Management and Information*, 1(1), 37-69.
- Vanhaecht, K., de Witte, K., Depreitere, R., & Sermeus, W. (2006). Clinical pathway audit tools: a systematic review. *Journal of Nursing Management*, 14(7), 529–537.
- Wade, D., & de Jong, B. (2000). Recent advances in rehabilitation. *British Medical Journal*, 320(7246), 1385.
- Welch, P. (1981). On the problem of the initial transient in steady-state simulation. *IBM Watson Research Center*.