Chapter 5
Applying the Cardinality–Constrained Approach in Health Care Systems: The Home Care Example

Ettore Lanzarone and Giuliana Carello

Abstract Many approaches are applied to deal with uncertainty in health care optimization problems. However, a recently proposed technique, namely, the cardinality–constrained approach, is only marginally applied in health care. This approach accounts for a given degree of uncertainty with a reasonable computational effort, providing a trade-off between computational time and robustness. In this paper, we apply such approach to the nurse-to-patient assignment problem under continuity of care arising in home care services. A linear programming model is developed for solving the problem, and the robustness is included in the formulation according to the cardinality–constrained approach. The overall robust model is applied to a Home Care provider operating in Italy, in order to evaluate its capability of reducing the costs related to nurses' overtimes, and to compare the results both with the real practice of the analyzed provider and with previously developed approaches. Relevant benefits are achieved by applying the proposed model in the practice, and results suggest that such benefits could be also achieved in other optimization problems within the health care domain.

5.1 Introduction

Uncertainty is a key feature of many health care optimization problems, which cannot be neglected and may have a significant impact on the problem solution. In locating emergency vehicles, uncertainty is associated to the availability

E. Lanzarone (✉)
Istituto di Matematica Applicata e Tecnologie Informatiche (IMATI), Consiglio Nazionale delle Ricerche (CNR), Via Bassini 15, 20133 Milan, Italy
e-mail: ettore.lanzarone@cnr.it

G. Carello
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, P.za Leonardo da Vinci 32, 20133 Milan, Italy
e-mail: giuliana.carello@polimi.it

A. Matta et al. (eds.), *Proceedings of the International Conference on Health Care Systems Engineering*, Springer Proceedings in Mathematics & Statistics 61,
DOI 10.1007/978-3-319-01848-5_5,

of ambulances [1], while in planning and scheduling operating room theaters uncertainty is mainly related to the duration of surgery [2]. Uncertainty also occurs in managing home care (HC) services, where sudden variations in the amount of service required by patients, which is in general highly variable, are the most critical and frequent random events.

Different approaches are usually applied to deal with uncertainty in optimization problems, such as probabilistic models, stochastic optimization approaches, and, more recently, the cardinality–constrained approach proposed in [3]. This accounts for a certain degree of uncertainty (which can be tuned) with a reasonable computational effort, providing a trade-off between needed computational time and robustness.

Although the approach seems to match many health care optimization problems, to the best of our knowledge it has not been often applied in this field so far (only four papers with keyword *health care* that cite [3] were found in March 2013 through a search on ISI web of knowledge and Scopus).

In this paper, we present an application of the cardinality–constrained approach to the nurse-to-patient assignment problem under continuity of care in HC. The approach can be easily applied to the problem and proved to produce good quality solutions with a reasonable computational effort. Therefore, it is also worthy of being tested on other health care optimization problems.

5.1.1 Home Care Service

HC consists of delivering medical, paramedical and social services to patients at their domicile rather than in hospital. This leads to a significant improvement in the quality of life for patients, as they continue to live at their home, and to considerable cost savings for the entire health care system, as hospitalization costs are avoided. Moreover, HC is a relevant and growing sector in western countries, due to the population aging, to the increasing of chronic pathologies, to the introduction of innovative technologies, and to the pressure of governments to contain health care costs.

Many resources are involved in delivering HC services, including nurses, other operators, support staff and material resources. In addition, the presence of peculiar constraints, such as the continuity of care and the operator risk of incurring burnout, makes the HC resource planning different from the planning problems arising in other health care systems.

Continuity of care means that a HC provider assigns only one nurse to each patient, called *reference nurse*, and the assignments are kept for a long period. This is an important quality indicator since patients are always cared for by the same nurse, instead of continuously developing new relationships, and potential loss of information among operators is avoided. However, continuity of care limits the flexibility of the service, and some providers do not adopt it to increase the operational efficiency. In general, for a good balance between quality and flexibility, the continuity of care should be preserved at least for critical patients (e.g., palliative patients) or patients with particular needs.

5.1.2 Literature Review

The literature about HC management can be mainly divided into two groups: the first one deals with daily schedule of visits and routing of nurses, and the second one deals with staff planning and management from a mid-term and long-term perspective. The nurse-to-patient assignment is related to the mid-term management. Different features may be considered, such as the continuity of care and the uncertainty in patients' demands.

Nurse-to-patient assignment has been rarely studied as a stand alone problem (i.e., not considering the scheduling [4]) and, to the best of our knowledge, the assignment problem taking into account the continuity of care is only marginally addressed in the literature [5–7]. Besides, continuity of care is often considered as an objective rather than a strict requirement [8]. If continuity of care is not considered, the assignment problem turns out to be an assignment of operators to visits rather than to patients, in which the aim is to jointly optimize the operator-to-visit assignment and the scheduling and routing problem [9, 10]. In districts with a limited territorial extension (e.g., in Europe), the impact of travel times on scheduling and routing is not very significant; hence, assignment and scheduling problems are separately solved since the joint optimization requires a significant computational effort and, consequently, reduces the length of the considered time horizon.

As mentioned, uncertainty inherently arises in HC due to unpredictable changes in patients' needs. In [11] it is managed by representing the whole system as a Markov chain and developing admittance policies for patients.

The nurse-to-patients assignment problem, in which both continuity of care and demand uncertainty are considered, has been rarely addressed in the literature. The problem was tackled with stochastic programming [6] and with analytical policies [7]. However, both these approaches proved limited even if they improve the quality of the assignment with respect to those actually applied by the HC structures. The stochastic programming approach is based on scenario generation and, due to the high number of patients and the associated demand variability, requires to include a very high number of scenarios. Only a limited number of them can be consequently considered for a computationally acceptable solution. Therefore, a high expected value of perfect information (EVPI) and a low value of the stochastic solution (VSS) are obtained [6]. The analytical policies are related to strict assumptions regarding, e.g., the shape of workload probability density functions, the number of assignable patients, and the number of periods in the planning horizon [7].

With the cardinality–constrained approach, we aim at exploiting the potentialities of a linear programming model rather than an analytical approach, without the necessity of generating scenarios.

5.2 Robust Assignment Model

We consider the problem of assigning a set of patients P to a set of nurses I over a time horizon T divided into a set of time slots. Three continuity of care requirements are considered:

- *Hard continuity of care*: patients must be assigned to only one reference nurse for the entire time horizon. These patients are partitioned into two subsets P_c^a and P_c^n. Patients in P_c^a are already under treatment and assigned at the beginning of the time horizon, and they keep their assignment. Patients in P_c^n start their treatment at the beginning of the time horizon, and they are not yet assigned.
- *Partial continuity of care*: the reference nurse can be changed from time slot to time slot. However, each reassignment is penalized by a cost γ to keep the number of reassignments limited. As for the previous case, these patients are partitioned into two subsets P_{pc}^a and P_{pc}^n. Patients in P_{pc}^a are already under treatment at the beginning of the time horizon, while patients in P_{pc}^n start their treatment at the beginning of the time horizon.
- *No continuity of care*: patients can be assigned to more than one nurse even in the same time slot and the assignments can be changed from a time slot to another without penalties (set P_{nc}).

The division in districts is taken into account: a parameter m_{ij} is given for each nurse $i \in I$ and patient $j \in P$, which is equal to 1 if nurse i operates in the district of j, and 0 otherwise.

The amount of working time required by patient $j \in P$ in time slot $t \in T$ is an uncertain parameter r_{jt}, with expected value $\bar{r}_{jt}$ and maximum value $\bar{r}_{jt} + \hat{r}_{jt}$. Each nurse $i \in I$ has an amount of available working time per time slot v_i, and overtime must be paid if v_i is exceeded. The overtime cost depends on its amount. A set of overtime levels L_i are defined for each nurse $i \in I$, and two parameters are given for each level $l \in L_i$: a threshold Δ_i^l and a cost per time unit c_l for each overtime unit above $v_i + \sum_{k=1}^{l-1} \Delta_i^k$ and below $v_i + \sum_{k=1}^{l} \Delta_i^k$.

The problem consists of assigning all of the patients to the nurses, according to the required continuity of care, with the aim of minimizing the overtime costs and the number of reassignments for patients with partial continuity of care.

The problem is modeled as follows. A binary variable x_{ji} is defined for each patient $j \in P_c^a \cup P_c^n$ and nurse $i \in I$ ($x_{ji} = 1$ if j is assigned to i during the whole time horizon, and 0 otherwise). Similarly, a binary variable ξ_{ji}^t is defined for each patient $j \in P_{pc}^a \cup P_{pc}^n$, nurse $i \in I$ and time slot $t \in T$ ($\xi_{ji}^t = 1$ if nurse i is in charge of patient j during time slot t, and 0 otherwise). The assignments of patients to reference nurses before the considered time horizon are described with parameters $\tilde{x}_{ji}$ ($\tilde{x}_{ji} = 1$ if $j \in P$ is initially assigned to $i \in I$, and 0 otherwise). Furthermore, a binary variable y_j^t is introduced for each patient $j \in P_{pc}^a \cup P_{pc}^n$ ($y_j^t = 1$ if the assignment of patient j is changed from time slot $t-1$ to time slot t, and 0 otherwise). Finally, the fraction of time needed by $j \in P_{nc}$ in time slot $t \in T$ provided by nurse $i \in I$ is represented by a continuous variable $\chi_{ji}^t \in [0,1]$. The overtime assigned to each nurse $i \in I$ in time slot $t \in T$ is described by a continuous variable w_{it}^l for each level $l \in L_i$, which represents the extra workload related to c_l.

The objective function aims at minimizing the overtime costs and the number of reassignments: these two parts are both relevant, as the first one reduces the burnout risk, while the second one guarantees a suitable quality of provided service.

$$\min \left\{ \sum_{i\in I}\sum_{t\in T}\sum_{l\in L_i} \left(c_l w_{it}^l \right) + \gamma \sum_{j\in P_{pc}^a \cup P_{pc}^n} \sum_{t\in T} y_j^t \right\} \tag{5.1}$$

subject to:

$$\sum_{i\in I} m_{ij} x_{ji} = 1, \qquad \forall j \in P_c^a \cup P_c^n \tag{5.2}$$

$$\sum_{i\in I} m_{ij} \xi_{ji}^t = 1, \qquad \forall j \in P_{pc}^a \cup P_{pc}^n, t \in T \tag{5.3}$$

$$\sum_{i\in I} m_{ij} \chi_{ji}^t = 1, \qquad \forall j \in P_{nc}, t \in T \tag{5.4}$$

$$\sum_{j\in P_c^a \cup P_c^n} r_{jt} x_{ji} + \sum_{j\in P_{pc}^a \cup P_{pc}^n} r_{jt} \xi_{ji}^t + \sum_{j\in P_{nc}} r_{jt} \chi_{ji}^t \le v_i + \sum_{l\in L_i} w_{it}^l, \quad \forall i \in I, t \in T \tag{5.5}$$

$$0 \le w_{it}^l \le \Delta_i^l, \qquad \forall i \in I, t \in T, l \in L_i \tag{5.6}$$

$$x_{ji} \ge \tilde{x}_{ji}, \qquad \forall i \in I, j \in P_c^a \tag{5.7}$$

$$y_j^t \ge \xi_{ji}^t - \xi_{ji}^{t-1}, \qquad \forall t \in T \setminus \{t_1\}, j \in P_{pc}^a \cup P_{pc}^n, i \in I \tag{5.8}$$

$$y_j^{t_1} \ge \xi_{ji}^{t_1} - \tilde{x}_{ji}, \qquad \forall j \in P_{pc}^a, i \in I \tag{5.9}$$

Constraints (5.2)–(5.4) guarantee that each patient is assigned to a suitable nurse; constraints (5.5) compute nurse workloads and overtimes for each level; constraints (5.6) set the thresholds for the overtime workload; constraints (5.7) guarantee that patients in P_c^a do not change their assignment at the beginning of the time horizon; constraints (5.8) and (5.9) compute the number of reassignments.

To deal with uncertainty in constraints (5.5) we apply the cardinality–constrained robust model proposed in [3]. The basic idea of the approach is that only a subset of the uncertain parameters are likely to assume their maximum value simultaneously. The approach provides solutions which are feasible even if at most Γ uncertain parameters assume their worst possible value (i.e., the maximum value) rather than their expected value. As the solution must be feasible for any choice of Γ parameters for each constraint, the subset which represents the worst possible case is selected. The impact is then computed exploiting duality properties, yielding to a linear formulation.

We apply the cardinality–constrained approach to the proposed formulation by including, for each nurse and time slot, three subsets S_c^{it}, S_{pc}^{it} and S_{nc}^{it} of patients assigned to i (with $S_c^{it} \subseteq P_c^a \cup P_c^n$, $S_{pc}^{it} \subseteq P_{pc}^a \cup P_{pc}^n$, and $S_{nc}^{it} \subseteq P_{nc}$), whose demand charged to nurse i in time slot t is equal to the maximum treatment time $\bar{r}_{jt} + \hat{r}_{jt}$. Cardinality is constrained as at most Γ_c^i, Γ_{pc}^i and Γ_{nc}^i patients (with Γ_c^i, Γ_{pc}^i and Γ_{nc}^i integer) are assumed to belong to these subsets, respectively. The charged demand of all other patients is the expected value $\bar{r}_{jt}$.

The robustness is taken into account considering the worst possible charge for each nurse i at each time slot t in constraints (5.5). As example, for patients requiring hard continuity of care, the term $\sum_{j \in P_c^a \cup P_c^n} r_{jt} x_{ji}$ is replaced with:

$$\sum_{j \in P_c^a \cup P_c^n} \bar{r}_{jt} x_{ji} + \max_{\substack{S_c^{it} | S_c^{it} \subseteq P_c^a \cup P_c^n, \\ |S_c^{it}| = \Gamma_c^i}} \left\{ \sum_{j \in S_c^{it}} \hat{r}_{jt} x_{ji} \right\}$$

Let us denote the maximum related to a given solution $\{x^*\}$ with $\beta_c^{it}(x^*, \Gamma_c^i, t)$:

$$\beta_c^{it}(x^*, \Gamma_c^i, t) = \max_{\substack{S_c^{it} | S_c^{it} \subseteq P_c^a \cup P_c^n, \\ |S_c^{it}| = \Gamma_c^i}} \left\{ \sum_{j \in S_c^{it}} \hat{r}_{jt} x_{ji}^* \right\}$$

This is computed for each nurse i and time slot t by solving the following linear programming problem:

$$(\mathscr{P}_c^{\beta it}) = \max \sum_{j \in P_c^a \cup P_c^n} \hat{r}_{jt} x_{ji}^* z_{ji}^t \tag{5.10}$$

$$\sum_{j \in P_c^a \cup P_c^n} z_{ji}^t \leq \Gamma_c^i \tag{5.11}$$

$$0 \leq z_{ji}^t \leq 1, \quad \forall j \in P_c^a \cup P_c^n \tag{5.12}$$

where $z_{ji}^t \in [0,1]$ are continuous variables which represent the choice of the elements in subset S_c^{it}. The associated dual problem is:

$$(\mathscr{D}_c^{\beta it}) = \min \sum_{j \in P_c^a \cup P_c^n} \pi_{jit}^c + \Gamma_c^i \zeta_{it}^c \tag{5.13}$$

$$\zeta_{it}^c + \pi_{jit}^c \geq \hat{r}_{jt} x_{ji}^*, \quad \forall j \in P_c^a \cup P_c^n \tag{5.14}$$

$$\pi_{jit}^c \geq 0, \quad \forall j \in P_c^a \cup P_c^n \tag{5.15}$$

$$\zeta_{it}^c \geq 0 \tag{5.16}$$

where ζ_{it}^{c} are the dual variables associated with (5.11), and π_{jit}^{c} the dual variables associated with $z_{ji}^{t} \leq 1$ (5.12).

Optimal values $(\mathscr{P}_c^{\beta it})$ and $(\mathscr{D}_c^{\beta it})$ coincide and, therefore, the maximum can be replaced by $\sum_{j \in P_c^a \cup P_c^n} \pi_{jit}^{c} + \Gamma_c^i \zeta_{it}^{c}$ adding the following variables and constraints to the model:

$$\zeta_{it}^{c} + \pi_{jit}^{c} \geq \hat{r}_{jt} x_{ji}, \quad \forall i \in I, j \in P_c^a \cup P_c^n, t \in T$$

$$\zeta_{it}^{c} \geq 0, \quad \forall i \in I, t \in T$$

$$\pi_{jit}^{c} \geq 0, \quad \forall i \in I, j \in P_c^a \cup P_c^n, t \in T$$

The same idea is applied to $\sum_{j \in P_{pc}^a \cup P_{pc}^n} r_{jt} \xi_{ji}^{t}$ and $\sum_{j \in P_{nc}} r_{jt} \chi_{ji}^{t}$, thus obtaining the robust cardinality–constrained version of the model.

In this way, each feasible solution remains feasible if any subset of at most Γ_c^i, Γ_{pc}^i and Γ_{nc}^i patients, respectively, require the highest number of visits.

5.3 Real Case Analysis

Computational tests are run in order to evaluate the applicability of the proposed approach to a real HC provider. The quality of the solutions and their impact when applied to realistic scenarios are taken into account.

The analysis is conducted on the same HC provider already studied in other papers dealing with assignment techniques under continuity of care [6, 7], so as to compare the outcomes of the proposed model with other approaches. Furthermore, a patient stochastic model to estimate the future patients' demands is available for this provider [12]. The considered HC provider operates in the north of Italy, covering a region of about 800 km^2, with about 1,000 patients assisted at the same time by about 50 nurses. The provider includes three independent divisions, and the analysis is carried out for the nurses of the largest one. The division consists of six districts and the analysis is carried out in four of them where more than one nurse is present (Table 5.1). The assignments are planned considering the districts as independent.

Table 5.1 Analyzed districts

Name of the district	Code of territory	Skill of the nurses	Number of nurses
NPA	A	Non-palliative	8
PA	A	Palliative	3
NPB	B	Non-palliative	4
NPC	C	Non-palliative	5

Table 5.2 Analyzed instances

Type of continuity for palliative	Γ values	Robust solution	Non-robust solution
C	1	Conf. A	Conf. E
	2	Conf. B	Conf. E
random 80% C	1	Conf. C	Conf. F
and 20 % PC	2	Conf. D	Conf. F

5.3.1 Experimental Setup

We consider data related to 26 weeks from April to September 2008 [6, 7]. The model is applied according to a rolling approach, and each time slot t is 1 week. An initial assignment of nurses is computed at the initial week (named week 0) considering all patients as newly admitted ones, while the successive assignments are provided on a rolling basis: at the beginning of each week, newly admitted patients are included in the mix and discharged patients are excluded. For each rolling week, the planning horizon includes the considered week and the next seven ones ($T = 8$). The assignments computed for the first horizon week are then kept, and the model is solved again for the next rolling week taking into account the information about patients assigned in the previous rolling weeks. This is consistent with the policy of the analyzed HC service provider, where assignments are mainly decided at the beginning of each week on a weekly basis. The initialization at week 0 is obtained neglecting the robustness (i.e., all patients require the expected demand $\bar{r}_{jt}$).

The reassignment penalty γ is assumed equal to 2.5, and 10 overtime levels are considered ($l = 1,\ldots,10$), with $c_l = l \;\forall l$ and $\Delta_i^l = 0.1 v_i \;\forall i,l$.

The number of patients in charge at each week and their features are taken from the historical data of the provider (considering real arrivals of new patients and real discharges), while patients' demands are estimated with the stochastic model proposed in [12]. The expected demand $\bar{r}_{jt}$ and the maximum demand $\bar{r}_{jt} + \hat{r}_{jt}$ of each patient are taken from an empirical probability density function given by such stochastic model (maximum value $\bar{r}_{jt} + \hat{r}_{jt}$ is taken neglecting the right tail of the distribution with probability 0.1).

The continuity of care requirement for each patient is determined based on his/her characteristics. Patients belong to 15 different care profiles (CPs) [12] and the type of continuity of care required by each patient is once decided according to the CP when the patient is first considered. For non-palliative patients, low intensity CPs require no continuity of care, middle intensity CPs partial continuity of care, and high intensity CPs hard continuity of care. Two different configurations are taken into account for palliative patients: either they all require hard continuity of care, or they require hard or partial continuity according to a random choice: each palliative patient is randomly considered requiring hard continuity of care (with probability 0.8) or partial continuity (with probability 0.2).

Two levels of robustness are considered, either $\Gamma_c^i = \Gamma_{pc}^i = \Gamma_{nc}^i = 1, \forall i$ or $\Gamma_c^i = \Gamma_{pc}^i = \Gamma_{nc}^i = 2, \forall i$. Moreover, also the case in which the robustness is neglected is studied (Table 5.2).

5.4 Results

The model has been implemented with OPL 5.1 and solved with CPLEX; computational tests have been run on a PC equipped with CPU Intel Core i7 1.73 GHz and 6 GB of RAM. A stopping condition on the gap is set so as to limit the computational time (1 % for configurations A and C; 4 % for configurations B and D). No stopping condition is set for the non-robust configurations E and F.

Table 5.3 shows the computational time, the objective function and the number of reassignments for patients with partial continuity of care. Results are expressed in terms of minimum, maximum and average values among the weeks from 1 to 25; week 0 is excluded as it refers to the non-robust initialization.

Results show that, with the adopted gaps, computational times are reasonable for any configuration. The objective function increases with the values of Γ_c^i, Γ_{pc}^i and Γ_{nc}^i due to both the overtime costs and the number of reassignments, as the demand of the worst scenario increases and more robust solutions are selected. The overtime cost is significantly affected by the degree of robustness of the solution, as the maximum demands of patients belonging to S_c^{it}, S_{pc}^{it} and S_{nc}^{it} have an impact on the overall workload.

Then, the question arises on how a robust solution behaves if no patients require the maximum amount of care. For evaluating the behavior of the solutions with respect to the expected demands, the assignments are applied assuming that each patient is requiring the expected demand $\bar{r}_{jt}$. The obtained overtime costs are reported in Table 5.4 in terms of minimum, maximum and average values among the weeks from 1 to 25.

It can be seen that robustness determines an increase of overtime costs. However, it is worth noting that, when considering the expected demands, the robust assignment is not significantly penalized with respect to the optimal non-robust counterpart. Indeed, overtime expected costs are always lower than the double of the non-robust case.

Table 5.3 Computational time in seconds, objective function and number of reassignments

	Computation time			Objective function			Num. of reassignments		
Configuration	Min	Max	Average	Min	Max	Average	Min	Max	Average
A	5	115	31	59.3	157.1	98.4	0	4	0.8
B	4	7,339	505	177.7	484.5	299.6	0	8	1.2
C	5	987	83	101.6	265.6	168.2	0	4	0.8
D	4	7,580	644	181.7	579.9	349.0	0	6	1.2
E	1	3	2	3.9	27.8	13.6	0	1	0.1
F	1	2	2	3.9	27.8	13.4	0	1	0.0

Table 5.4 Overtime cost from the objective function and overtime cost recomputed with the expected demands

	Overtime cost			Overtime expected cost		
Configuration	Min	Max	Average	Min	Max	Average
A	59.3	149.6	96.4	7.9	35.6	16.5
B	175.2	482.0	296.6	6.3	43.4	20.5
C	99.1	265.6	166.2	7.5	42.1	20.3
D	179.2	564.9	346.0	8.3	41.6	20.9
E	3.9	27.8	13.4	3.9	27.8	13.4
F	3.9	27.8	13.3	3.9	27.8	13.3

Table 5.5 Executed mean overtime cost per nurse: minimum, maximum and average values among the weeks from 1 to 25

	Sample paths			Real execution		
Configuration	Min	Max	Average	Min	Max	Average
A	0.00	9.90	1.85	0.25	16.27	4.55
B	0.00	9.24	1.80	0.07	13.29	4.55
C	0.00	10.98	1.51	0.10	13.59	3.72
D	0.00	9.46	1.55	0.41	11.53	3.95
E	0.00	13.35	2.17	0.64	16.12	5.81
F	0.01	11.79	2.30	0.92	19.16	5.94

5.4.1 Execution of the Assignments

Each obtained solution is applied to 10 sample paths (generated with the same procedure of [6, 7]) and to the real historical patients' demands.

The quality of the solutions is analyzed in terms of the mean overtime cost per nurse. This is obtained at each week as the ratio between the total cost of the district (computed with the same levels $c_l = l$ and thresholds $\Delta_i^l = 0.1 v_i$) and the number of nurses in the district. This indicator is directly taken for the execution with the historical demands, while for the sample paths the analyzed indicator is the average at each week among the paths. Hence, for each configuration and district, the result is the list of average costs over the weeks in two cases: executed with the historical demands or averaged among the sample paths (Table 5.5). We remark that planned costs, reported in Table 5.4, refer to the entire planning horizon (8 weeks), while for the execution only the first week of the planning horizon is extracted from each rolling week.

Results show that robust solutions perform better than their non-robust counterparts, both for sample paths and real data; thus, robustness provides the desired cost savings. To give an idea of the obtained cost savings, we can assume that one unit of cost corresponds to about 15 euros. Considering that the 4 districts include 20 nurses (see Table 5.1) and that the observed period refers to 25 weeks, each cost reduction of 1 unit corresponds to a global saving of 7,500 euros in the period. As example, comparing solution C with the corresponding non robust solution F, a global cost saving of 16,650 euros is observed for the real execution, and of 5,925 euros for the average among the paths.

Considering the detail of each district, the main benefits are obtained in districts NPA and PA both in terms of average and maximum values. A low benefit is observed in district NPC and hardly any benefit in district NPB. Then, it seems that larger benefits are obtained in the presence of critical patients with higher demands (i.e., palliative patients) or many nurses.

It must be stressed that non robust models are always solved to optimality, while an optimality gap is accepted for the robust counterparts. A robust, even if sub-optimal, solution computed in reasonable time is able to improve the solution upon its optimal non-robust counterpart on the considered case study.

Finally, if compared to other methodologies applied to this instance [6, 7], the cardinality–constrained approach is able to solve problem in a lower computational time (while including the stochasticity with the scenario generation of the stochastic programming approach requires huge computational times) with few assumptions on the demands (while the analytical approach based on stochastic ordering requires to introduce many assumptions on the shape of the density functions).

5.5 Discussions and Conclusions

In this paper, we apply the robust cardinality–constrained approach proposed in [3] in the health care area and, in particular, to the nurse-to-patient assignment in HC services under continuity of care. HC is chosen because of its novelty within the health care domain and the high randomness related to the workload amount, which is strongly higher than in other services. Thanks to this approach, the deterministic assignment model is easily modified to take into account the uncertainty in patients' demands, without the necessity of assuming probability density functions or deriving a relevant number of stochastic scenarios.

The proposed model has been tested on a set of generated instances and on historical data, and it provides good quality solutions in terms of overtime costs. The application of the cardinality–constrained approach to HC is then promising. Moreover, due to the general characteristics of HC within the health care domain, the obtained benefits could extend to other health care problems.

The main limit of the proposed approach is that patients are not allowed to have a demand for visits lower than the expected value $\bar{r}_{jt}$, while in the real practice some patients have a demand lower than the expected value. Such limit could be overcome by introducing different levels of demand for each patient rather than the two ones considered in this work; this will be the aim of our future work.

References

1. Brotcorne, L., Laporte, G., Semet, F.: Ambulance location and relocation models. Eur. J. Oper. Res. **147**, 451–463 (2003)
2. Denton, B.T., Miller, A.J.: Optimal allocation of surgery blocks to operating rooms under uncertainty. Oper. Res. **58**, 802–816 (2010)
3. Bertsimas, D., Sim, M.: The price of robustness. Oper. Res. **52**(1), 35–53 (2004)
4. Boldy, D., Howell, N.: The geographical allocation of community care resources – a case study. J. Oper. Res. Soc. **31**, 123–129 (1980)
5. Hertz, A., Lahrichi,N.: A patient assignment algorithm for home care services. J. Oper. Res. Soc. **60**(4), 481–495 (2009)
6. Lanzarone, E., Matta, A., Sahin, E.: Operations management applied to home care services: the problem of assigning human resources to patients. IEEE Trans. Syst. Man Cybern. A **42**(6), 1346–1363 (2012)
7. Lanzarone, E., Matta, A.: A cost assignment policy for home care patients. Flex. Serv. Manuf. J. **24**(4), 465–495 (2012)
8. Nickel, S., Schroder, M., Steeg, J.: Mid-term and short-term planning support for home health care services. Eur. J. Oper. Res. **219**, 574–587 (2012)
9. Trautsamwieser, A., Hirsch, P.: Optimization of daily scheduling for home health care services. J. Appl. Oper. Res. **3**(3), 124–136 (2011)
10. Rasmussen, M.S., Justesen, T., Dohn, A., Larsen, J.: The home care crew scheduling problem: preference–based visit clustering and temporal dependencies. Eur. J. Oper. Res. **219**, 598–610 (2012)
11. Koeleman, P.M., Bhulai, S., Van Meeresbergen, M.: Optimal patient and personnel scheduling policies for care-at-home service facilities. Eur. J. Oper. Res. **219**, 557–563 (2012)
12. Lanzarone, E., Matta, A., Scaccabarozzi, G.: A patient stochastic model to support human resource panning in home care. Prod. Plan. Control **21**(1), 3–25 (2010)