Operations Management Applied to Home Care Services: The Problem of Assigning Human Resources to Patients

Ettore Lanzarone, Andrea Matta, and Evren Sahin

***Abstract*—In recent years, home care (HC) service systems have been developed as alternatives to conventional hospitalization. Many resources are involved in delivering HC service, including different categories of human resources, support staff, and material resources. One of the main issues encountered while planning human HC resources is the patient assignment problem, i.e., deciding which operator(s) will take care of which admitted patient given some sets of constraints (e.g., the continuity of care). This paper addresses the resource assignment problem for HC systems. A set of mathematical programming models to balance the workloads of the operators within specific categories are proposed. The models consider several peculiarities of HC services, such as the continuity of care constraint, operators' skills, and the geographical areas which patients and operators belong to. Given the high variability of patient demands, models are developed under the assumption that patients' demands are either deterministic or stochastic. The analysis of the results obtained from a real case study demonstrates the applicability of the proposed models as well as the benefits that stem from applying them. Moreover, the obtained results show that an acceptable level of continuity of care cannot be obtained without modeling the continuity of care as a hard constraint. The analysis under continuity of care also shows the high value of information and the difficulties of fully balancing workloads with the application of standard techniques.**

***Index Terms*—Continuity of care, home care (HC), mathematical programming, operations management, patient assignment, stochastic patient demand.**

I. Introduction

THIS PAPER deals with a well-known type of service system, namely home health care or more simply home care (HC), which has been developed in recent years in developed countries as an alternative to conventional hospitalization. HC service aims at providing medical, paramedical, and social services to patients at their own domicile. This service has been developed in a number of countries to decrease patients' hospitalization rates, improve their quality of life, and reduce costs across the entire health care system [1], [2]. Many resources are involved in HC service delivery, including different categories of operators (e.g., nurses, physicians, physiotherapists, social assistants, psychologists), support staff, and material resources (either for care or transportation activities in a typically vast territory covered by the HC service provider).

Each patient receives care from a set of operators that belong to different categories. While a patient is always assigned to the nursing service, he/she may also require the involvement of other categories of operators. Patients admitted to an HC system are divided into classes, such as palliative care and non-palliative care patients. Within each class, patients are again categorized into several care profiles (CP) based on the specific needs of their treatment in terms of human and material resource requirements and the associated costs.

Operators of each category are assigned to patients who have a CP for which the operator has the appropriate skills. Moreover, in large HC service providers, operators are divided into territorial groups and take care (only or preferably, based on the policy of the provider) of patients who reside within their territory [1], [2]. Therefore, each operator belongs to a district, which is defined according to his/her skills and the territory where he/she can work. The assignment problem refers to the decision regarding which operator(s) among the compatible ones will take care of which patients.

Many random events affect the service delivery. The main randomness stems from unexpected changes in patient conditions, expressed as variations in the number, frequency, and duration of visits, which makes the activity of the HC service provider highly uncertain. Other unexpected events are changes in patient visiting hours, operator and material resource unavailabilities, scheduling problems, and random events perturbing the transportation of human and material resources.

Many of the HC service providers pursue the continuity of care [3]–[5]. Continuity of care means that a patient is assigned to only one operator of each category, named the reference or the principal operator, who follows the entire patient's care pathway during his/her sojourn at the HC structure and preferably provides all of the visits pertinent to his/her category. This is considered to be an important indicator of HC service quality, both because information loss among operators is avoided and because the patient receives care from the same person rather than having to continuously develop new relationships with new operators [3], [4]. However, to increase operational efficiency, some other HC service providers do not adopt the concept of continuity of care. In this case, the care service to the patient is provided by any appropriate operator who has sufficient capacity in the considered period.

Manuscript received February 23, 2011; revised October 1, 2011, March 26, 2012, and June 25, 2012; accepted July 9, 2012. Date of current version October 12, 2012. This paper was recommended by Associate Editor M. P. Fanti.

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

A. Matta is with the Dipartimento di Meccanica, Politecnico di Milano, 20133 Milan, Italy (e-mail: andrea.matta@polimi.it).

E. Sahin is with the Laboratoire Génie Industriel, Ecole Centrale Paris, 92295 Châtenay-Malabry Cedex, France (e-mail: evren.sahin@ecp.fr).

Digital Object Identifier 10.1109/TSMCA.2012.2210207

Appropriate resource planning is crucial for avoiding process inefficiencies and overloaded operators, which may lead to issues such as the burnout syndrome [6]. In practice, most HC service providers lack methodologies and tools suitable for managing the complex logistic and organizational activities that have to be coordinated to support efficient care delivery [5], [7]. Resources are generally planned on the basis of the standard treatments associated with the current patients' CPs. This information is neither customized according to patients' specificities within the CP nor frequently updated to follow the dynamics of their needs. As a consequence, the feasibility of plans is mined by the daily occurrences of disruptive events, and plans are frequently modified to the detriment of the continuity of care. Moreover, despite the importance of planning HC human resources, there is a lack of studies dealing with the assignment problem in the literature. Particularly, as presented in Section II, only few works consider the continuity of care issue, and none of them take into account patient demand variability.

The aim of this paper is to propose different decision-making models for the assignment problem that represent several real situations. For this purpose, a set of mathematical programming models have been formulated with the primary goal of balancing the workloads of the operators of a specific category while taking into account or neglecting the continuity of care. While operators' availabilities are considered deterministic in the models, patients' demands are assumed to be either deterministic or stochastic because of their high variability. In particular, the inclusion of the high variability of patient demand in resource planning for obtaining more efficient and robust plans is investigated, while in the literature this aspect is not exploited yet. Finally, the models are applied to the real case of a representative HC service provider to validate the proposed modeling approach.

This paper is organized as follows. In Section II, we survey the literature related to human resource planning in the HC context. In Sections III–VI, we develop different mathematical formulations of the HC assignment problem, and in Section VII we map these formulations to real organizations. The description and the results of computational experiments carried out on a real case are reported in Sections VIII and IX. Finally, Section X points out some concluding remarks and directions for future research.

II. Human Resource Planning in Home Care Services

Despite the growing importance of HC services in practice, the number of investigations dealing with operations management problems within the HC context is still modest in comparison with earlier models developed for hospitals [2].

Basically, the literature on human resource planning in HC deals with four main issues: dimensioning human resources, partitioning a territory into a given number of districts (i.e., the districting problem), assigning operators to visits (or to patients), and defining the schedule of visits and the routing of operators. Assignment and scheduling/routing represent the most important topics in the existing investigations.

A. Human Resources Dimensioning

This problem deals with determining the number of operators with particular skills necessary to meet the predicted care demand with satisfactory service quality at minimum cost.

In this context, De Angelis [8] studied the resource dimensioning problem for HC organizations addressed to AIDS patients. This research evaluates the suitability of budgets assigned to HC activities by public-health authorities in the city of Rome, Italy. The author proposed a stochastic linear programming model which is linked to an epidemiological model and integrates the uncertainty in the number of admitted patients and the levels of care required by patient classes. This model aims at maximizing the number of newly admitted patients based on resource availability constraints, a minimum standard of service, variability of demand, transition rates among classes, and budget constraints.

The funding problem in the context of HC was also studied by Busby and Carter [9], who created a decision tool able to determine the tradeoffs between the cost, the patient waiting time, and the service quality, which is defined as the number of visits per patient. Based on these tradeoffs, the HC service provider and the government are able to negotiate the funding levels necessary to satisfy the demand as well as the associated waiting time.

B. Districting Problem

This problem consists of appropriately grouping operators and patients into clusters, known as districts, according to relevant criteria (e.g., territory and skill versus patient's condition compatibility). Hence, operators belonging to a given district have the same skills and work in the same territory.

Blais *et al.* [10] proposed a multi-criteria approach to partition a region into districts. Criteria related to operators' mobility and workload balancing are combined into a single objective function, while criteria related to the indivisibility of basic units, the respect of borough boundaries and connectivity are considered to be hard constraints. The problem is solved using the Tabu search technique. Benzarti *et al.* [11] developed two formulations of the districting problem by separately considering different patient profiles. These two mathematical formulations consider the compactness and care workload balance criteria to be either hard constraints or objective functions. Lahrichi *et al.* [12] reviewed the optimality of Blais *et al.* [10] and proposed two solutions to alleviate the risk of overloading certain districts. The first solution is a dynamic approach based on a regular updating of districting, while the second solution involves the splitting of operators into two groups: one group of operators assigned to a fixed district and another group of flexible operators who are allowed to work in all or a part of the whole territory.

C. Human Resource Assignment to Visits or to Patients

Once the districting is defined, human resources associated with each district have to be assigned to patients in an equitable manner.

Three different complexity factors can be included or excluded while defining the assignments: the joint consideration of the assignment and the scheduling and routing of visits, the inclusion of the continuity of care constraint, and the stochasticity of patient demand. In the literature, the assignment problem has not been solved by including all of these factors, due to the high complexity of the problem and the computational time needed to consider them simultaneously.

With regard to the joint consideration of assignment and scheduling-routing, this problem is not relevant in HC settings where districting is not modified during the assignments and districts have a limited territorial extension. Indeed, in this case, the travel time can be assumed to be independent from the specific sequence of visits and, therefore, the assignment problem is not affected by the scheduling and routing of visits. This is the typical case of several European HC service providers, while in the U.S. and in Canada providers usually consist of districts with large territorial extensions. In this case, the assignments should be determined together with the scheduling and routing processes.

To our knowledge, only one work deals with the assignment problem separated from scheduling. Boldy and Howell [13] developed an approach for allocating HC resources that involves two steps: the assessment process, in which the nature and level of services required per patient are evaluated, and the allocation procedure, which aims at equitably distributing service units among the districts. This is a case study type work based mainly on survey information (related to patients, to services already available within the territory, and to the amount of HC actually received), which aims at deciding for an appropriate geographical split of a care givers. However, the aim of this work is not the development of a planning model, and it is not strictly related to the assignment of operators.

With regard to the continuity of care, if this constraint is not required and patients can be cared for by several operators, the assignment problem consists of assigning human resources to visits rather than to individual patients. Hence, the assignment planning and the routing problems are the same. As a matter of fact, the literature without continuity of care deals with the joint determination of assignments and schedules (see Section II-D).

If the continuity of care is required, operators are assigned to patients as reference operators and the assignments can be determined either jointly or separately from the scheduling and the routing. In the literature, Borsani *et al.* [5] coupled a simple assignment model and a scheduling model: the output of the assignment is the input to the scheduling. The objective of the assignment is to ensure workload balance among operators while respecting some constraints such as the continuity of care, qualification requirements, and geographical coherence constraints. Hertz and Lahrichi [14] proposed two mixed integer programming models for allocating operators to patients. One model includes linear constraints and a quadratic objective function, while the other includes nonlinear constraints and is solved by a Tabu search heuristic. The objective is to balance the nurses' workloads by minimizing a weighted sum of the visit load (based on the weight of each visit), the case load (due to the number of patients assigned) and the travel load (related to the distances traveled) while respecting constraints related to maximum acceptable loads and continuity of care. The possibility of assigning a patient to a nurse who does not work within the patient's district to reduce workload imbalances is also considered.

The present paper focuses on the assignment in the presence of districts with a limited territorial extension, which are defined before the assignment problem is solved and not modified. In comparison with the existing works, we aim at giving a more complete modeling framework of the possible variants of the assignment problem in HC under continuity of care. Moreover, different from the literature, we analyze the effects of patient demand variability on the assignment by comparing the solutions obtained in a deterministic or a stochastic demand context.

D. Scheduling and Routing

The scheduling and routing problem without continuity of care (i.e., without an explicit assignment of patients to operators) has been investigated by several authors. In this context, the aim is to associate a set of visits to each operator without considering the patients related to these visits. Begur *et al.* [15] developed a module for the daily scheduling of operators' activities, which simultaneously assigns visits to operators and generates the sequence in which visits have to be provided. This is based on a heuristic approach that combines a set of procedures (e.g., k-optimal procedure, sweep algorithm, insertion procedures) to build daily operators' routes. The objective is to minimize the total travel time while respecting constraints related to route construction, operators' time windows, and skills requirements. Cheng and Rich [16] addressed the daily scheduling problem as a multi-depot vehicle routing problem with time windows and compatibility information. The objective is to minimize the total cost associated with the overtime hours assigned to full-time nurses and the hours given to part-time nurses. The problem is formulated as a mixed integer linear programming model and solved by two-phase heuristics. Eveborn *et al.* [17] developed a decision support system, named Laps Care, in which the scheduling problem is formulated as a set partitioning model and solved by a repeated matching algorithm. The objective is to minimize a cost related to travel time, scheduled hours, preferences, etc., while respecting criteria such as time windows for visits, operators' skill requirements, and accomplishment of each visit by one operator. Bertels and Fahle [18] proposed a combination of linear programming, constraint programming, and heuristics to assign operators to visits and to optimally sort the visits assigned to each operator. The goal is to minimize the total transportation cost and maximize the satisfaction of both patients and operators, while considering a variety of soft constraints (e.g., patient-operator affinities and preferences for certain visits). Thomsen [19] formulated the daily scheduling problem as a vehicle routing problem. The objective is to minimize the total travel time, the number of shared visits (carried out by two operators) and unlocked visits (carried out by non-reference operators), while respecting the patients' and operators' time windows, assigning at least one visit to each operator and beginning and ending shared visits at the same time. This problem is solved by an insertion heuristics and the Tabu search technique. Akjiratikarl *et al.* [20] addressed the daily scheduling problem as a vehicle routing problem with time windows and solved it with the particle swarm optimization meta-heuristics. The objective consists of minimizing the total distance traveled while respecting constraints related to patients' and operators' time windows and assignment of each visit to only one operator. Chahed *et al.* [21] considered the scheduling problem for the chemotherapy at home, where the production and distribution problems are dependent because the anti-cancer drugs produced at hospital and distributed to patients' homes have limited lifetimes. The authors identify six models based on three criteria: the respect of patients' time windows, the objective function considered (to minimize the distribution cost or to maximize the earnings achieved per patient visited), and the distribution of drugs by HC personnel or by an external logistic service. More recently, Ben Bachouch *et al.* [22] developed a mixed linear programming model at scheduling level which aims at minimizing the total distance traveled by nurses. This model considers several constraints, including patients' and nurses' time windows, nurses' meal breaks, continuity of care, maximum distance between two consecutive visits by the same nurse, and each nurse's route beginning and ending at the HC facility.

One common characteristic of these scheduling and routing models is that none of them consider the demand uncertainty. Moreover, it can be seen that the assignment problem under continuity of care is usually solved together with the scheduling and routing.

III. Definition of the Assignment Problem

The patient assignment problem in HC service requires deciding which operator(s) of a certain category will deliver the care service (i.e., the visits) to which patients during a certain planning horizon.

Operators are divided into districts, based on their main skills and working territories. Each district groups homogeneous operators in terms of skill, territory, and, consequently, intervention times. In our setting, districts are defined before solving the assignment problem, and the assignment-type decisions do not consider other operational decisions such as scheduling or routing of visits.

Since the scheduling and routing of visits are not included, the interaction among the different categories of operators is not explicitly considered in the proposed models and does not affect the solution of the assignment problem. Hence, the assignments are defined separately for each category of operators, and the models examine only one category at a time. This approach reflects the real condition of HC service providers having districts of limited territorial extensions, which are not modified during the assignments, where the different categories of operators are managed and assigned independently. The interactions among the different categories will be considered during the scheduling of visits when particular sequences among visits (e.g., precedence, simultaneity) have to be respected.

Resources are assigned to balance the operators' workloads while simultaneously satisfying a set of constraints specific to HC operations. The assignment problem is solved either at a fixed frequency (e.g., day or week) or when a certain amount of newly admitted patients are reached. The choice of the assignment updating procedure depends on the organization of the HC service provider and the number of patients under care.

TABLE I
MODELS

	One reference operator	Multi operator assignment
One-district model (I)	Objective: Eq.1 Constraints: Eq.2-5	Objective: Eq.1 Constraints: Eq.4,5,13-20
N-districts model (II)	Objective: Eq.7 Constraints: Eq.2,3,5,6,8	Objective: Eq.7 Constraints: Eq.5,6,8,13-20
N-districts model with penalties (III)	Objective: Eq.12 Constraints: Eq.2,5,6,9-11	Objective: Eq.12 Constraints: Eq.5,6,10,11,13-19,21

Mathematical programming models with the corresponding equations for the deterministic configuration. Eq.22-Eq.26 are added in the case of multiple operator assignment if the presence of a primary operator is required.

Operators are assigned to patients at the beginning of a planning horizon composed of several periods k (with $k = 1, \ldots, n_w$). Assignments are made on a rolling time basis: each time the problem is solved, a new planning horizon n_w is considered and the assignment problem is solved over this planning horizon. This logic, similar to a rolling horizon-based production planning process, enables the integration of future information about the evolution of the system.

The assignment problem varies depending on the type of service offered, the organization of the HC service provider and the data available during the decision-making process. These variants enable to combine several aspects of the problem such as the repartition of districts, the flexibility of resources, and the continuity of care.

The following sections describe different models for addressing relevant variants of the problem. Models with continuity of care and deterministic patient demands are presented in Section IV, Section V analyzes the continuity of care relaxation, while the inclusion of stochastic patient demand is described in Section VI. A summary of the developed formulations is provided in Table I.

The developed models are different from the classical literature, due to the possibility of assigning visits to operators in excess of his/her capacity and the inclusion of the continuity of care constraint [23], [24].

IV. Deterministic Models Under Continuity of Care

In these models, each newly admitted patient i (with $i = 1, \ldots, n_{pat}$) has a deterministic demand for care $d_{i,k}$ from the examined category of operators for each period k of the planning horizon in which he/she is in charge. This demand $d_{i,k}$ is estimated from the data available at the assignment moment and is expressed as a non-negative continuous variable modeling the time for visits (expressed in hours) needed by patient i in period k until the estimated discharging period. Consequently, if a patient is estimated to leave the HC service in any period k^* prior to n_w, then $d_{i,k} = 0$ for $k > k^*$. This time also includes the travel to reach the patient's home. For any patient to operator assignment within the district, the travel time is assumed to be independent of the specific assignment and the sequence of visits, according to the considered limited territorial extension of districts.

Each operator j of the available operators (with $j = 1, \ldots, n_{op}$) is characterized by three aspects:

– The set of patient classes he/she can handle, referred to as *skill* in the remainder of the paper. In the case of complex care providers, each operator has a main skill and additional skills. The main skill refers to the class of patients for which the operator is best suited to care (full knowledge of patient characteristics, intervention times equal to the standard, etc.). Additional skills enable an operator to handle patients belonging to a different category than his/her main skill, but the intervention times are expected to be longer in these cases.
– The geographical area where he/she operates. This criterion is considered in cases in which HC service providers are partitioned into several territorial areas, where each area groups the operators that work in the same zone.
– His/her capacity $a_{j,k}$ in each considered period k, which is the total amount of time (expressed in hours) that the operator can work according to his/her working contract. The capacity is assumed to be deterministic.

Continuity of care implies that only one operator is assigned to each patient and that it is not possible to change a previously established assignment. As a consequence, at the time the assignment problem is solved, only newly admitted patients must be assigned to one operator, while each previously admitted patient retains his/her assigned operator. Hence, each operator j has a certain time for visits $w^0_{j,k}$ in each period k of the planning horizon to provide to patients already assigned to him/her (whose demands are estimated and updated with the data available each time the assignment problem is solved) which is a non-negative continuous parameter already determined before the considered assignments.

A. Model I: One-District Model

The model described in this section (*Model I*) considers that operators and patients belong to the same district and includes the hard constraint of continuity of care. All of the operators are similar in terms of experience, and they each work with patients who require care associated with their main skill. Hence, the time spent on a visit is assumed to be independent of the operator who provides the visit. Furthermore, under continuity of care, only one reference operator of each category delivers services to the patient.

The utilization rate of operator j in period k of the planning horizon is used as the operator's workload indicator and is defined as the ratio $w_{j,k}$ to $a_{j,k}$. The utilization rate is usually lower than 1 so as not to saturate the operator, but it can take higher values under specific cases (e.g., the operator does overtime work, not included in $a_{j,k}$, for extra pay to satisfy a surplus demand). Our model does not explicitly consider the cost due to possible overtime work.

The objective of the assignment is to balance the operators' utilization rates. The measure of the balancing adopted in this work is the range between the maximum and minimum rates $w_{j,k}/a_{j,k}$ over the planning horizon. Specifically, the objective can be reached through a number of different formulations. To reduce the computational effort, the balancing is performed by maximizing the minimum utilization rate in each period k, meaning that any newly admitted patients are assigned to the operators with the lowest utilization. Hence, the objective function is to maximize the sum of h_k over the planning horizon, where h_k is the minimum utilization rate $w_{j,k}/a_{j,k}$ over all operators in period k, as follows:

$$maximize \sum_{k=1}^{n_w} h_k \tag{1}$$

$$s.t.$$

$$\sum_{j=1}^{n_{op}} x_{i,j} = 1 \qquad \forall i \tag{2}$$

$$w_{j,k} = w^0_{j,k} + \sum_{i=1}^{n_{pat}} d_{i,k} \cdot x_{i,j} \qquad \forall j, k \tag{3}$$

$$h_k \leq \frac{w_{j,k}}{a_{j,k}} \qquad \forall j, k \tag{4}$$

$$x_{i,j} \in \mathbb{N} \quad x_{i,j} \geq 0 \qquad \forall j, k. \tag{5}$$

The model decision variables are:

– assignments of newly admitted patients $x_{i,j}$ (Boolean variables with $x_{i,j} = 1$ if the newly admitted patient i is assigned to operator j; 0 otherwise);
– total time for visits $w_{j,k}$ supplied by operator j in period k, including newly admitted patients and previously assigned patients.

Equation (2) implies that all newly admitted patients must be assigned to only one operator. Equation (3) computes the total workload of each operator j in period k. Inequality (4) expresses the minimum utilization rate h_k, which is maximized in the objective function. Equation (5) ensures the non-negativity and the integrality of $x_{i,j}$.

Usually, in these kinds of models, the objective function is the minimization of the maximum utilization rate. However, in this case, some of the utilization of each operator is fixed due to the previously assigned workload and patients, and it is not possible to reduce the maximum workload below this fixed amount.

This model can be applied independently as many times as the number of different categories of operators to manage. In the case of the assignment of a *team* of operators of different categories all together (e.g., for therapy administration), the model should be extended to consider all of the categories together by introducing a minimum utilization rate specific for category (to reach different balancings) or a unique h_k among all the categories (to pursue a global balancing).

B. Model II: N-Districts Model With Independent Districts

In a more general context, the HC service provider is divided into a certain number of districts, and each operator belongs to a district according to his/her skills and the territory which he/she works in.

In this configuration, the number of districts n_g, the membership of an operator to a district, and the compatibility between each operator and each patient are assumed to be known before the assignment problem is solved. In particular, patients and operators are partitioned into sets Ω_g (with $g = 1, \ldots, n_g$), where Ω_g includes patients and operators belonging to district g. The compatibility $c_{i,g}$ between the newly admitted patient i and the operators belonging to district g is defined based on territorial compatibility and operator skills versus patient care needs. All of the operators associated with a given district g are equal in terms of the time they spend to visit a certain patient (as in *Model I*); therefore, $c_{i,g}$ does not depend on operator j in district g.

Districts may be managed independently as stand-alone districts or in a coordinated way to compensate for local demand fluctuations among the different districts. Depending on the values assigned to these compatibilities, the assignment problem can be solved for an HC service provider which works either with independent or integrated districts.

Resource management in independent districts is quite simple because each district is considered an autonomous decision center. *Model II* describes this condition, in which a patient can be assigned to only one district.

In this case, the compatibilities $c_{i,g}$ are Boolean parameters assumed to be as follows: $c_{i,g} = 1$ if patient i can be assigned to an operator of district g and 0 otherwise; only one $c_{i,g}$ for each patient i is equal to 1 to maintain the separation among districts. With respect to *Model I*, the following compatibility constraint is added in *Model II* to ensure that $x_{i,j} \geq 0$ only if $c_{i,g} = 1$ and $x_{i,j} = 0$ elsewhere:

$$x_{i,j} \leq c_{i,g} \qquad \forall i, j \quad g : j \in \Omega_g. \tag{6}$$

The minimum utilization rate among the operators has to be maximized within each district, resulting in the following objective function:

$$maximize \sum_{g=1}^{n_g} \sum_{k=1}^{n_w} h_{k,g} \tag{7}$$

where the index g refers to district g as follows:

$$h_{k,g} \leq \frac{w_{j,k}}{a_{j,k}} \qquad \forall j, k \quad g : j \in \Omega_g. \tag{8}$$

The other constraints are the same as those reported in (2), (3) and (5).

Model I can be considered to be a special case of *Model II* with $n_g = 1$ and $c_{i,g} = 1 \ \forall i, g$.

C. Model III: N-Districts Model With Integrated Districts

The HC service provider could prefer to coordinate operators among the districts (e.g., to fully exploit some special skills). In this case, time performance is a key variable to be considered because of the geographical distribution of patients and the skills of operators. Indeed, because one rationale for the district composition is the territorial repartition of patients and operators, the travel time included in the visit duration may differ if the patient is assigned to operators that work in different districts. Moreover, operators could be slower when they deliver care using their additional skills rather than their main skill.

Model I and *Model II* assume that all operators assignable to a certain patient provide visits of the same duration. Hence, they are assumed to have the same time performance. To deal with variable service times, *Model III* is proposed.

This model includes the possibility to assign a patient to an operator who does not belong to the patient's district (meaning either that the operator does not work in the same geographical area where the patient lives or that the operator provides care using his/her additional skills), with the penalty of increasing the effective operator engagement over the standard time for visits provided in-district.

To allow these out-of-district assignments, the compatibilities $c_{i,g}$ are modified with respect to *Model II*. They are defined as positive continuous parameters with $c_{i,g} = 1$ if patient i can be assigned to district g without a penalty, $c_{i,g} = 0$ if i cannot be assigned to g and $c_{i,g} > 1$ if i can be assigned to g with a penalty. The difference $(c_{i,g} - 1)$ is a positive real number representing the penalty factor for the visit.

When a patient is assigned to an operator who works in a different geographical area, the penalty factor is based on the estimated extra time spent by an operator of district g to reach the territory of patient i and come back. A more accurate evaluation of the penalties can be achieved in providers where the assignment is solved together with the scheduling and routing of visits (not considered in this paper). In this case, the penalty can be reduced in the presence of more than one patient in the same territory not belonging to the operator's district, to take into account the possibility of sequencing these out-of-district visits. When a patient of a given category is assigned to an operator who has to use his/her additional skills, the penalty factor is determined by the difference between the time spent by an operator who occasionally takes care of the considered patient category and that spent by an operator who usually cares for this category.

The assignment of an operator to an out-of-district patient leads to an increase in the time spent for the visits. Therefore, the workload computation of *Model I* and *Model II* is modified as follows:

$$w_{j,k} = w_{j,k}^0 + \sum_{i=1}^{n_{pat}} d_{i,k} \cdot c_{i,g} \cdot x_{i,j} \qquad \forall j, k \quad g : j \in \Omega_g \tag{9}$$

where each $w_{j,k}^0$ that operator j provides in period k to his/her previously assigned patients includes the penalty factors related to previously decided out-of-district assignments.

Moreover, the objective function is also modified, because the maximization of $h_{k,g}$ used to reduce the computational effort could push all patients to be assigned with a penalty with no control of workload imbalances among the districts. In this case, the objective is to minimize the utilization rates that are higher than the mean utilization rate of the corresponding district. For each operator j and period k, the over utilization rate $y_{j,k}$ is the difference between the utilization rate of j in k and the mean utilization rate of the district in k as follows:

$$y_{j,k} \geq \frac{w_{j,k}}{a_{j,k}} - \frac{\sum_{j\in\Omega_g} w^0_{j,k} + \sum_{i\in\Omega_g} d_{i,k}}{\sum_{j\in\Omega_g} a_{j,k}} \quad \forall j, k\; g : j \in \Omega_g \quad (10)$$

$$y_{j,k} \geq 0 \qquad \forall j, k. \quad (11)$$

The mean utilization is obtained with the demands $d_{i,k}$ of all new patients of the district, without including any penalties that may result from their assignments. In this way, the number of assignments with penalties is limited, and out-of-district assignment only occurs if some operators have utilization rates higher than the average. Reducing the total overutilization means reducing any imbalances among the operators. Therefore, the objective function minimizes $y_{j,k}$ for each operator j and each period k, i.e., the sum of these positive quantities is as follows:

$$minimize \sum_{j=1}^{n_{op}} \sum_{k=1}^{n_w} y_{j,k}. \quad (12)$$

Alternatively, only one mean utilization can be considered for the entire provider, to include the reduction of the imbalances among the districts.

As for *Model I*, it is possible to extend the mathematical formulation of *Model II* and *III* to assign a reference team of operators belonging to different categories.

V. Continuity of Care Relaxation

In the case of complete continuity of care, only one operator is assigned to each patient and the assignments do not change while the patient remains in the HC system. This constraint can be relaxed to obtain partial continuity of care or completely neglected.

In particular, the continuity of care can be relaxed in two ways: first, more than one reference operator of a category can be assigned to a patient (multi-operator assignment); second, a change in the reference operator of a patient can be allowed over two consecutive periods. If both of these aspects are relaxed, a complete absence of continuity of care is obtained. Such variants are analyzed in the following sections.

A. Multi-Operator Assignment

In the case of multiple reference operators of a certain category, the minimum number n_i and the maximum number N_i of reference operators assignable to each patient i are introduced. These parameters can be differentiated based on the specific patient, or they can be assumed to be equal to n and N $\forall i$. Hence, for *Model I* and *II*, (2) is modified as follows:

$$n_i \leq \sum_{j=1}^{n_{op}} x_{i,j} \leq N_i \qquad \forall i. \quad (13)$$

Moreover, the decision variable $d_{i,k,j}$, which corresponds to the time for visits required by patient i in period k and allocated to operator j, is introduced to split each patient's demand among his/her reference operators. While the total time for visits $d_{i,k}$ required by patient i in period k is a parameter in the model, $d_{i,k,j}$ is a non-negative continuous decision variable which is subject to three further constraints as follows:

$$\sum_{j=1}^{n_{op}} d_{i,k,j} = d_{i,k} \quad \forall i, k \quad (14)$$

$$d_{i,k,j} \leq d_{i,k} \cdot x_{i,j} \quad \forall i, j, k \quad (15)$$

$$d_{i,k,j} \geq x_{i,j} \cdot \alpha_i \cdot d_{i,k} \quad \forall i, j, k. \quad (16)$$

Equation (14) splits parameter $d_{i,k}$ among the operators, (15) constrains that only the reference operators are considered in the splitting, and (16) ensures that each reference operator has a minimum amount of the patient's demand in each period of the planning horizon.

The minimum amount α_i related to patient i (as a relative fraction of the entire patient's demand) is a non-negative parameter assumed before the assignment, defined so as to obtain a feasible solution. This minimum amount follows the real practice of HC service providers, where a reference operator provides at least a certain number of visits to his/her assigned patients.

With regard to the previously assigned patients, they maintain their reference operators, but the splitting of their demands is reconsidered when new assignments are provided. Hence, the initial workload $w^0_{j,k}$ of operator j in period k of the planning horizon is detailed as the sum of the demands of his/her already assigned patients. For this reason, the demand for care $\delta_{l,k}$ of each already assigned patient l (with $l = 1, \ldots, n_{charged}$) in each period k is included in the models. Similar to $d_{i,k}$, parameter $\delta_{l,k}$ is expressed in hours and estimated from the data available at the moment when the assignment problem is solved. $\delta_{l,k}$ is then split among the reference operators of patient l adopting the decision variable $\delta_{l,k,j}$, which is the time for visits required by patient l in period k and assigned to operator j. $\delta_{l,k,j}$ is subject to the same constraints of $d_{i,k,j}$ as follows:

$$\sum_{j=1}^{n_{op}} \delta_{l,k,j} = \delta_{l,k} \qquad \forall l, k \quad (17)$$

$$\delta_{l,k,j} \leq \delta_{l,k} \cdot z_{l,j} \qquad \forall l, j, k \quad (18)$$

$$\delta_{l,k,j} \geq z_{l,j} \cdot \alpha_l \cdot \delta_{l,k} \qquad \forall l, j, k \quad (19)$$

where $z_{l,j}$ is a Boolean parameter (already decided before the considered assignment with $z_{l,j} = 1$ if the already admitted patient l is assigned to operator j and $z_{l,j} = 0$ otherwise) and α_l is a non-negative parameter analogous to α_i.

In this way, for *Model I* and *II*, the total workload of each operator j in period k is determined as follows:

$$w_{j,k} = \sum_{l=1}^{n_{charged}} \delta_{l,k,j} + \sum_{i=1}^{n_{pat}} d_{i,k,j} \qquad \forall j, k \quad (20)$$

where $x_{i,j}$ and $z_{l,j}$ do not appear because $d_{i,k,j} = 0$ if $x_{i,j} = 0$ and $\delta_{i,k,j} = 0$ if $z_{i,j} = 0$.

For *Model III*, (2) and (9) change into (13)–(19) and (21)

$$w_{j,k} = \sum_{l=1}^{n_{charged}} d_{l,k,j} \cdot \gamma_{l,g} + \sum_{i=1}^{n_{pat}} d_{i,k,j} \cdot c_{i,g} \quad \forall j,k \quad g : j \in \Omega_g \tag{21}$$

where the compatibility $\gamma_{j,g}$ between patient j to district g is analogous to $c_{i,g}$.

B. Multi-Assignment With Primary Operator

In the presence of more than one reference operator, it can be required that one of the reference operators assigned to patient i is his/her *primary* operator, who provides a significant amount of the visits required by patient i during the care pathway. In other words, the time for visits provided by the primary operator is required not to be lower than an imposed threshold t_i. This is expressed as a relative fraction of patient demand, which is obviously greater than α_i (i.e., $\alpha_i < t_i < 1\ \forall i$).

A new series of constraints are written in this case, introducing a Boolean parameter $q^*_{l,j}$ which is determined before the considered assignments and a Boolean decision variable $q_{i,j}$ as follows:

$$d_{i,k,j} \geq q_{i,j} \cdot t_i \cdot d_{i,k} \qquad \forall i,j,k \tag{22}$$

$$\delta_{l,k,j} \geq q^*_{l,j} \cdot t_l \cdot \delta_{l,k} \qquad \forall l,j,k \tag{23}$$

$$\sum_{j=1}^{n_{op}} q_{i,j} = 1 \qquad \forall i \tag{24}$$

$$q_{i,j} \leq x_{i,j} \qquad \forall i,j \tag{25}$$

$$q_{i,j} \in \mathbb{N} \quad q_{i,j} \geq 0 \qquad \forall j,k. \tag{26}$$

$q_{i,j} = 1$ and $q^*_{l,j} = 1$ if j is the primary operator; $q_{i,j} = 0$ and $q^*_{l,j} = 0$ otherwise. Equations (22) and (23) force $d_{i,k,j}$ and $\delta_{l,k,j}$ to be greater than the threshold if $q_{i,j} = 1$ and $q^*_{l,j} = 1$, respectively. Equations (24) and (25) impose that $q_{i,j} = 1$ for only one operator among the reference ones. Equation (26) ensures the non-negativity and the integrality of $q_{i,j}$.

C. No Continuity Between Two Consecutive Periods

In the case that no continuity of operator is required between two consecutive periods, all the proposed models can be adopted. In this condition, they are solved with each patient treated as a newly admitted patient. Hence, $w^0_{j,k} = 0\ \forall j,\ k$ or $n_{charged} = 0$ depending on the model.

Moreover, $n_w = 1$ because new assignments are made at each period, and it is not useful to balance the workload over more than one period.

D. Complete Absence of Continuity of Care

Under complete absence of continuity of care, the maximum number of operators assignable to a patient is set equal to the number of available operators, and the assigned operators are not preserved over time. In this case, the models are solved with each patient treated as a newly admitted patient ($w^0_{j,k} = 0\ \forall j,\ k$ or $n_{charged} = 0$), with $n_w = 1$, $n_i = 1$, and $N_i = n_{op}\ \forall i$ without any threshold t_i.

VI. Stochastic Patient Demand

The models described so far allow for solving the assignment problem in only one scenario of patient demands, where this single scenario usually consists of the expected values of the demands. In this way, the information about demand variability is not used, the obtained assignments are optimized for the specific scenario, and they lack of robustness. On the contrary, more robust plans could be obtained while solving the assignment problem in more than one scenario of future patient demands at the same time.

For this purpose, each one of the proposed models is modified to include a number n_s of scenarios for patient demands. In this configuration, each newly admitted patient i has a demand for care $d_{i,k,s}$ in period k and scenario s (with $s = 1, \ldots, n_s$) and, in the case of partial continuity, each already admitted patient l has a demand for care $\delta_{l,k,s}$ in period k and scenario s. Then, the workloads are evaluated in each scenario s ($w_{j,k,s}$ and $w^0_{j,k,s}$ are the total and initial workloads of operator j in period k and scenario s, respectively). On the contrary, assignments $x_{i,j}$ and compatibilities $c_{i,g}$ are kept constant because the same assignments have to be provided in all of the scenarios and patients belong to a district independent of their demands.

Under continuity of care, *Model I* is modified as follows:

$$maximize \sum_{k=1}^{n_w} \sum_{s=1}^{n_s} p_s \cdot h_{k,s} \tag{27}$$

$$s.t.$$

$$w_{j,k,s} = w^0_{j,k,s} + \sum_{i=1}^{n_{pat}} d_{i,k,s} \cdot x_{i,j} \qquad \forall j,k,s \tag{28}$$

$$h_{k,s} \leq \frac{w_{j,k,s}}{a_{j,k}} \qquad \forall j,k,s \tag{29}$$

$$\text{Eq.2, Eq.5}$$

where $h_{k,s}$ is the minimum of the utilization rates in period k and scenario s, while p_s is the probability of scenario s occurring. In this way, the objective is mediated among the scenarios, and a robust solution is searched for.

The expected workload $\hat{w}_{j,k}$ of operator j in period k is finally extracted as follows:

$$\hat{w}_{j,k} = \sum_{s=1}^{n_{sc}} p_s \cdot w_{j,k,s}. \tag{30}$$

Model II is modified introducing the minimum utilization rate $h_{k,g,s}$ over all operators of district g in period k and scenario s, and maximizing $\sum_{k=1}^{n_w} \sum_{g=1}^{n_g} \sum_{s=1}^{n_s} p_s \cdot h_{k,g,s}$.

Model III is modified including the scenario s in (9) and an overutilization $y_{i,k,s}$ in each scenario s. Hence, the objective function is to minimize $\sum_{j=1}^{n_{op}} \sum_{k=1}^{n_w} \sum_{s=1}^{n_s} p_s \cdot y_{j,k,s}$.

In the case of multi-operator assignment, the splitting of the demand has to be considered in each scenario s: $d_{i,k,j,s}$ and $\delta_{l,k,j,s}$ are the amount of $d_{i,k,s}$ and $\delta_{l,k,s}$ assigned to operator j in scenario s, respectively. The associated constraints in (14)–(19) and the total workload computation in (20) and (21) are then expressed for each scenario s.

Finally, in the presence of a primary operator, only one threshold t_i is considered over the scenarios, and (22) and (23) are expressed for each scenario s.

Also, the capacity $a_{j,k}$ of operator j in period k is a random variable in practice, but its variability is considerably lower than one of the patients' demands in real HC service providers. Hence, $a_{j,k}$ is assumed to be deterministic also in this setting, without reducing the robustness of the solution.

VII. Mapping Models to Real Organizations

In this section, the choice of the most appropriate assignment model for different HC service providers is discussed. Providers differ depending on their organizational structures [1], [2], [7]. Indeed, we can distinguish providers (usually linked to hospitals) that care for only specific categories of patients, providers that take care of high-profile critical patients, and large providers that care for all categories of HC patients. Typically, the first two types are private providers, while the third type consists of public providers that must guarantee service to all categories of patients in a certain geographical area.

The organizational structure has an impact on the districting scheme of the HC service provider, i.e., the territorial organization (one single territory, several independent or integrated territories) and the skills of operators (operators adequate for all patients or only for some of them). Moreover, public and private non-profit providers usually tend to preserve continuity of care, while for-profit providers do not always preserve continuity of care due to the cost related to this constraint.

In the mapping provided below, HC service providers are classified according to their territorial organization and then to the relationship between the skills of the operators and the classes of patients serviced.

A. Small Providers

Small providers cover a limited region and do not divide their territory into geographical areas. Three main typologies can be distinguished, according to the classes of cared patients:

- *Providers for only one class of patient*: all operators are usually equally skilled (main skill) for the particular class of patient under care. Each operator can take care of any patient, and the provider is not divided into districts at all. *Model I* can be used to solve the assignment problem in such organizations.
- *Providers with different classes of patients and focused team structures*: different classes of patients are present in the provider, and each operator is skilled for only one class, resulting in rigid division of operators. Patients are usually divided into palliative and non-palliative care, but the number of classes can be greater than two. The assignment problem is solved after the operators are districted according to their respective skills. Each district contains only operators taking care of a certain patient class, and operators belong to only one district. *Model II*, or *Model I* within each district, can be adopted to solve this assignment problem.
- *Providers with different classes of patients and flexible operators*: some of operators are skilled for different classes of patients, while others are skilled for only one class. The assignment problem is solved after grouping the operators with the same combination of main and additional skills into districts (i.e., two operators are in the same district if they can take care of the same classes of patients). A typical example is that of a provider caring for palliative and non-palliative patients: standard operators can only take care of non-palliative patients, while skilled operators can take care of both palliative and non-palliative patients. In this instance, a possible districting could be one group of operators for non-palliative patients, one for palliative patients and one of operators who can care for both types. *Model III* can then be adopted to solve the assignment problem, with compatibility $c_{i,g}$ of each patient i equal to 1 for all compatible districts. Finally, if some operators can take care of a patient class with a penalty (e.g., related to a longer visiting time than usual), the lack of full compatibility can be included in the coefficients $c_{i,g}$.

B. Large Providers With Independent Territorial Groups

Large providers serve a vast territory, which is usually divided into geographical areas to reduce the operators' travel times. The presence of independent geographical areas, where an operator cannot take care of patients residing outside of his/her territory, makes the assignment problem analogous to the case of small providers (Section VII-A). Hence, in each geographical area, the same possibilities of the small providers can be found. In the presence of independent districts within each geographical area, *Model I* can be applied to each district or *Model II* can be applied to the entire provider, also keeping the territorial groups separated by means of appropriate coefficients $c_{i,g}$. In the presence of integrated districts within one or more geographical areas, *Model III* is required.

C. Large Providers With Integrated Territorial Groups

In the presence of integrated geographical areas, operators can also take care of patients residing outside of their area. This is the case of large providers that face highly variable demand with operators who can work outside of their territory to compensate for a temporary overload of a district. Usually, a visit provided outside of one's territory implies some extra travel time. In this case, the corresponding single-territory model is extended to include the extra time in case of an out-of-territory assignment. Hence, *Model III* can be adopted. However, given the numerous complexities that arise from integrated resource management in large geographical areas, this case is not commonly encountered in practice.

VIII. Analysis of a Real Case

The assignment problem is solved in a real HC service provider in an effort to estimate the benefits resulting from the adoption of the proposed decision-making models. The

TABLE II
NURSES AND DISTRICTS

District	Territory	Skill of nurses	Number of nurses	Capacities
NPA	A	Non-Palliative	8	35, 40, 45, 50,50, 50, 50, 50
PA	A	Palliative	3	20, 30, 30
NPB	B	Non-Palliative	4	30, 35, 50, 50
PB	B	Palliative	1	35
NPC	C	Non-Palliative	5	30, 35, 40, 50, 50
PC	C	Palliative	1	35

Districts in the analyzed division of the provider, together the territory covered (A, B or C), the skill and the number of nurses, and the weekly capacity in hours for each nurse.

real case analysis is performed both under continuity of care and without continuity (i.e., multi-operator assignment and no continuity between consecutive periods). Under continuity of care, the variability of patients' demands is also considered so that to evaluate the robustness of the assignments. Finally, the solutions of the assignment problem obtained with the mathematical programming models are compared with the real practice of the provider.

A. Real Case Description

The analysis is carried out considering one of the largest Italian public HC providers. This provider includes three independent divisions with as many independent planning centers, and the analysis refers to the largest division. This provider is representative of a general class of HC facilities in terms of patient characteristics, organizational structure, and resource planning [1], [2]. Therefore, the obtained results can be considered to be general and extendable to several other HC service providers.

Patients are classified into two categories: palliative care patients and non-palliative care patients. Each newly admitted patient is assigned to only one reference operator for each category, and the assignment is preferably not changed to preserve the continuity of care. However, it may be changed to overcome the presence of idle or overloaded operators. The goal of the provider is to balance the workloads of the operators while preserving the continuity of care.

Each operator is characterized by his/her capacity, defined as the weekly time for visits according to his/her working contract (even if this value can be exceeded for extra pay under critical working conditions). Moreover, the different skills of operators (operators for palliative care and non-palliative care) and their territorial distributions are taken into account in the assignment. Three geographical areas are present in the analyzed division, and each operator cares for patients residing in his/her territory. Hence, the division consists of six districts (one for each combination of skill and territory), and the assignments are planned by considering the districts to be independent. However, an operator may sometimes be assigned to care for patients not belonging to his/her district. This out-of-district assignment is used to compensate for temporary workload imbalances among the districts, for infeasibilities of scheduling and routing or to account for changes in a patient's required care. For example, it is possible for a non-palliative patient to become palliative and, if the reference operator has adequate skills, the planner will not change the reference operator in such an instance due to the relationship already established between the patient and the non-palliative operator.

B. Experimental Setup

We consider the problem of assigning *reference nurses* in the largest division of the analyzed provider. Nurses are chosen because they provide the largest number of visits to HC patients, manage emergencies, meet short-term demand variations, and deal with highly uncertain workloads. Moreover, the continuity of care is usually pursued for nurses. Table II reports the skills and the capacities of the nurses of the analyzed division: only four districts have more than one nurse (i.e., NPA, PA, NPB, and NPC), and the analysis is provided for them.

Because of independent districts in the considered division, *Model II* is adopted, and the compatibilities $c_{i,g}$ are chosen to keep the districts strictly separated. The results are the same as those from adopting *Model I* in each district. The time period k is set to one week, and nurses are assumed to have constant capacity in every period (i.e., $a_{j,k} = a_j$).

The assignment problem is solved at a fixed frequency over a period of 26 weeks (from April to September 2008). An initial assignment is carried out for the initial week (week 0) for all of the patients currently in charge, while new patients are assigned on a rolling basis. Specifically, at the beginning of each week, the new patients admitted to the service in the previous week are assigned. This represents the policy of the analyzed HC service provider, where assignments are mainly decided at the beginning of each week.

The assignment problem is solved either by considering or neglecting the continuity of care.

Under perfect continuity of care, only one nurse is assigned to each patient, and this assignment is not changed in the following assignments. If a patient is discharged and readmitted during the 26 weeks, he/she maintains the previous assigned reference nurse at the readmission. The planning horizon of the assignments is set to eight weeks $(n_w = 8)$, because after two months the majority of critical patients with high demands leave the service. In this setting, the real numbers of patients and their classes over the considered weeks are taken, and patient demands are modeled either deterministic or stochastic to provide the analysis on the robustness of the assignment in the presence of the narrowing constraint of the continuity of care (Section VIII-E).

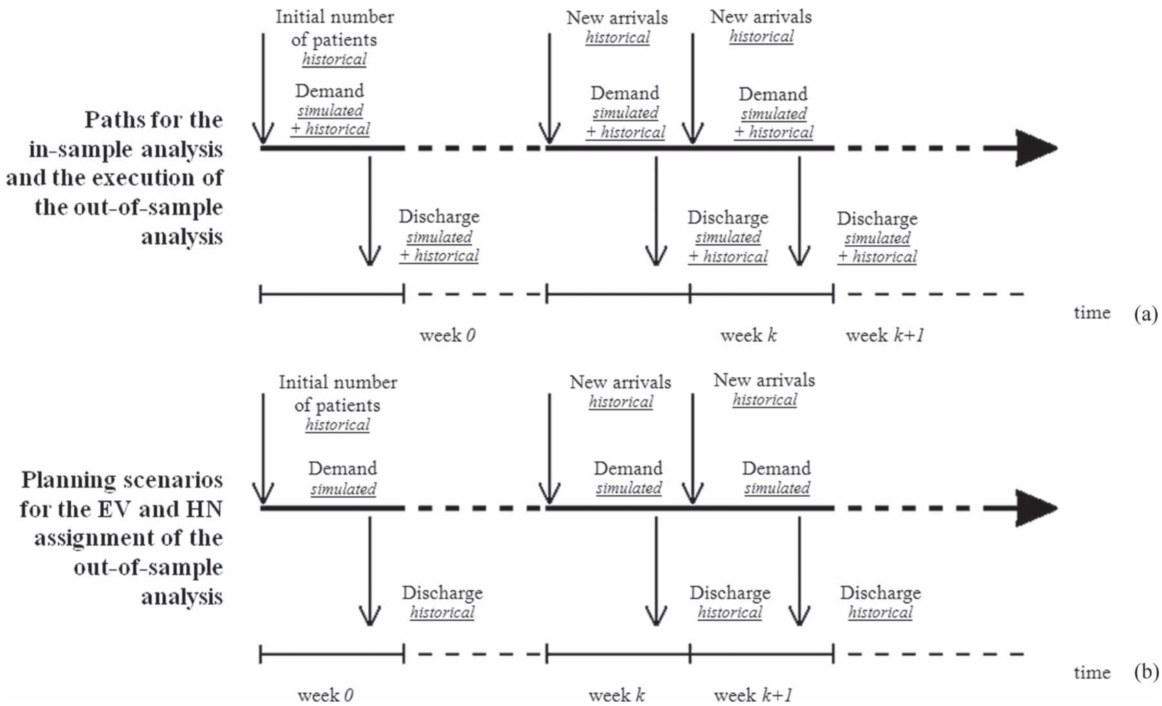

Fig. 1. Pattern for determining patient numerousness and demands in the analyses under continuity of care. (a) As for the n_p sample paths, the number of patients in charge and their demands are partially simulated: they are generated for the majority of patients (new arrivals are taken from historical data, while their care pathways, demands, and discharges are simulated) and taken from the historical data for long stay patients. (b) As for planning scenarios of EV and HN assignments in the out-of-sample analysis, the real number of patients in charge is taken from the historical data (replicating the initial number of patients, new arrivals and discharges) and their demands are simulated.

Two cases are analyzed without continuity of care:

- *multi-operator assignment*: the same values $n_i = N_i > 1$ are assumed $\forall i$, and the assigned reference nurses are kept stable over the future assignments. The planning horizon is set to eight weeks $(n_w = 8)$. Moreover, $\alpha_i = 0.15\ \forall i$ is set, and no primary nurse is taken into account.
- *no continuity of care between consecutive periods*: only one nurse is assigned to each patient at each week, without maintaining the assignment over time. In this case, the model is solved with $n_w = 1$ and all of the patients are treated as newly admitted patients.

Without continuity of care, the assignments are decided by taking into account the real numbers of patients and their classes over the considered weeks, while the patient demand $d_{i,k}$ of the planning horizon is modeled as deterministic and equal to the expected value estimated at the moment the problem is solved. The estimation of patients demand is based on the patient stochastic models reported in Lanzarone *et al.* [25]. This was built on the data of patients admitted in the analyzed HC provider since January 2004 and discharged before April 2008. In particular, the model provides estimates on future demands for each patient in charge in terms of empirical probability density functions, and the expected values are extracted from the corresponding density function.

The solutions of each experiment from week 1 to 25 provide information about the assignments and the *planned* workloads assigned to the nurses. All of the obtained assignments (under continuity of care or without continuity) are then evaluated either using the real observed values of patient demands (i.e., from historical data) or in a set of simulated paths, as described in Section VIII-C. These solutions, named *executed* in the following, provide information about the workloads that nurses would have faced given the assignments planned with the model.

Models were run using OPL 5.1 (ILOG IBM, Sunnyvale, CA, USA). The percent gap between the integer solution, and the best node was imposed to be lower than 0.5%.

Finally, the real workload that the nurses faced over the studied timeframe is considered. In practice, the provider assigned patients to operators trying to preserve the continuity of care, without adopting any assignment model and basing assignments only on ward sister's choices. Hence, this case (named *real* in the following) includes both real patient demands and real assignments decided in the HC organization.

C. Execution of the Assignments

The assignments are evaluated in two different configurations. They are executed in a set of n_p *simulated paths* or, alternatively, with the real observed values of patient demands from the historical data of the HC service provider (configuration named *historical* in the following).

The generation of simulated paths consists of a mix between a Monte Carlo simulation, from a semi-Markov model, and the real historical data [Fig. 1(a)]. The semi-Markov model, which is obtained modifying the approach proposed in the Markovian patient stochastic model of Lanzarone *et al.* [25], is used for the majority of patients. The real conditions of patients are taken at the beginning of the care pathway (or at week 0 for patients already in charge), and their requests and discharges are simulated for all of the future weeks starting from these real conditions. Specifically, for each patient, a path is generated by sampling the sojourn duration in a CP, the related demand for visits and the next CP in a Monte Carlo approach. Only the demands of long stay non-palliative patients are taken from their real historical values; these patients show a very low variability along with the time, they do not represent a relevant uncertainty source, and the variability of the entire mix of patients is not significantly reduced by this choice.

D. Performance Indicators

The mean utilization rate $\overline{u}_j$ of each operator j over the considered weeks is used as the workload level indicator for the operator. In the *planned* solutions, due to the rolling assignment approach, only the first week of the planning horizon gives information about the assigned workloads because the other weeks are influenced by future assignments. Then, the planned $\overline{u}_j$ is computed as the average of the utilization rates $w_{j,1}/a_j$ from week 1 to week 25 (or, in the stochastic case, a weighted value among the n_s scenarios). In each *execution*, the utilization rate in a week is computed with the weekly patient demand in the path (or, alternatively, with the real historical demand) and the provided assignments; then, $\overline{u}_j$ is computed as the average of these utilization rates.

The range of variation of $\overline{u}_j$ among all of the nurses belonging to a district, named Z and calculated as the difference between maximum and minimum $\overline{u}_j$ in the district, is assumed as the performance indicator of the assignment procedure in that district. In detail, the range of the planned utilizations provided by an assignment is referred to as $Z_{planned}$, while the range of each execution of a plan is referred to as $Z_{executed}$.

The range Z gives information about the workload balancing performance, i.e., a low Z value denotes well-balanced workloads. Moreover, better or worse workload balancing of a solution with respect to another is determined as the difference between the corresponding Z values.

In the absence of a perfect continuity of care, two indexes of the actual continuity over the considered weeks are defined. For each patient, the time for visits provided by the nurse who supplies the highest amount to the patient over all weeks is expressed as a fraction of the total patient's demand. To provide a continuity of care indicator for the entire division, this ratio is simply averaged over all patients (index λ_P) or calculated with an average weighted by each patient's total demand over the weeks (index λ_V).

E. Stochastic Programming Approach Under Continuity of Care

Under continuity of care, the variability of patients' demands is taken into account in the planning by means of a set of n_s generated *scenarios*.

The assignment problem is then solved under the expected value (EV), here and now (HN), and wait and see (WS) approaches [26]. In the EV approach, assignments are planned with the expected demands $d_{i,k}$ using the deterministic programming model. In the HN approach, assignments are planned taking into account the n_s scenarios all together with the stochastic programming model (in which probabilities p_s of scenario occurring are assumed the same for all of the scenarios, because of the very high number of existing scenarios and the finite number of included ones). In the WS approach, assignments are separately planned in each one of the scenarios using the deterministic programming model.

Each solution is executed in the n_p simulated paths obtaining a set of values $Z_{executed}$, and the performances are evaluated in terms of the $Z_{executed}$ expected value. Z_{eev} is the average of $Z_{executed}$ values given by the EV approach, Z_{hn} by the HN approach and Z_{ws} by the WS approach.

In a first case, the EV, HN, and WS solutions are obtained with the classical *in-sample analysis*, i.e., the same scenarios of patient demands are used both for planning and execution. The executions are performed in the simulated paths and each scenario for planning is extracted from a path considering the following 8 weeks (i.e., the planning horizon) starting from the current one in which the assignment is decided. Moreover, the expected demands for the EV solution are obtained as the average of patient's demands in the scenarios.

According to the stochastic programming theory [26], the three approaches would be compared in terms of the different values assumed by the objective function, computing the expected value of perfect information $(EVPI)$ and the value of stochastic solution (VSS). However, in this study, we are interested in the obtained ranges Z rather than in the values assumed by the objective function, because the maximization of the minimum utilization rate [(1), (7) and (27)] is introduced only to reduce the computational effort. Hence, we consider the same differences suggested by the stochastic programming theory, but applied to ranges Z. Thus, we adopt the indicators *modified EVPI (mEVPI)* and the *modified VSS (mVSS)*, which are defined according to the following equations:

$$mEVPI = Z_{hn} - Z_{ws} \tag{31}$$

$$mVSS = Z_{eev} - Z_{hn}. \tag{32}$$

Ideally, at each solving of the assignment problem, scenarios should include all of the possible combinations of weekly patients' demands all over the planning horizon. However, this results in a very high number of scenarios, while on the contrary a limited number has to be considered in order to obtain a computationally feasible solution.

This first *in-sample* analysis, even if unrealistic from an operational point of view, is performed with a limited number of scenarios to study the stochastic programming models without noises deriving from the impossibility to include the real execution among the scenarios.

A second *out-of-sample analysis* is also developed to consider the impossibility of computing all of the combinations by means of different scenarios for patient assigning with respect to workload evaluating. In other words, the assignments are evaluated out of the scenarios generated for making decisions.

The execution of plans and the WS planning are performed in the same n_p paths as in the in-sample analysis. Scenarios for the out-of-sample planning (planning for the EV and HN

TABLE III
NUMBER OF PATIENTS

District	Initial number of patients (week 0)	Admitted patients (weeks 1-25)	Total time for visits in hours (from week 1 to 25)												
			Historical data	Expected form patient stochastic model	Simulated paths										
					Average	# 1	# 2	# 3	# 4	# 5	# 6	# 7	# 8	# 9	# 10
NPA	230	186	8260	7363.48	7377.15	7544.75	7194.25	7672.50	7146.50	7535.00	7272.00	7546.75	7290.25	7576.75	6992.75
PA	33	33	1280	1283.10	1288.48	1171.00	1326.25	1479.00	1094.25	1068.25	1457.50	1285.50	1469.00	1307.00	1227.00
NPB	134	89	4609	4426.75	3944.98	3891.75	3953.00	4017.00	3837.00	3819.00	3963.75	3983.75	3985.00	4017.75	3981.75
PB	16	19	791	656.97	725.73	711.00	721.00	631.25	768.00	633.50	794.25	829.00	660.75	711.00	797.50
NPC	156	112	4739	4642.78	4709.25	4661.00	4724.50	4569.50	4858.25	4960.00	4733.75	4687.75	4583.75	4315.75	4998.25
PC	12	26	568	719.92	631.65	655.75	493.50	522.50	718.75	655.25	604.25	626.50	618.75	584.00	837.25
TOT	581	465	20247	19093.00	18677.24	18635.25	18412.50	18891.75	18422.75	18671.00	18825.50	18959.25	18607.50	18512.25	18834.50

Number of patients at week 0 and patients admitted from week 1 (starting on Monday April 6, 2008) to week 25 (ending on Sunday September 28, 2008) in each district together with the associated demands from nurses (total time for visits from week 1 to 25 expressed in hours). Values of the demands are in terms of real values from historical data, expected values from the patient stochastic model, and data generated in the paths.

approaches) are generated week by week as in the real practice of HC service providers [Fig. 1(b)]: in each week, the real mix of patients in charge is considered, and the future demand distributions for each patient in the planning horizon are obtained with the patient stochastic model of Lanzarone *et al.* [25]. Then, the expected demand is extracted (EV approach) from the distributions or scenarios are generated with a Monte Carlo technique (HN approach). The experiments are repeated for different values of n_s to evaluate the impact of the number of scenarios used in the HN approach.

In this analysis, the three solutions are compared again in terms of the different Z obtained. However, due to the different scenario generation for planning with respect to the paths for execution, it is not possible to refer them as $mEVPI$ and $mVSS$ [26]. For this reason, the differences are directly named $Z_{hn} - Z_{ws}$ and $Z_{eev} - Z_{hn}$.

In addition to the simulated paths, the assignments of the out-of-sample analysis are also executed with the real observed values of patient demands (from the historical data of the HC service provider). In this case, the WS approach is the clairvoyant solution, which is based on accurate estimates with no errors. This means to assume advance knowledge about the future demands of each patient. Even if it is unrealistic from an applicability point of view, this solution is informative to evaluate the best possible case given advance knowledge of the real patient demand.

IX. Results

A. Simulated Paths

The number n_p of simulated paths for the execution of plans is determined by imposing a half-width 95% confidence interval for Z_{ws} lower than 0.035 (smaller than each $Z_{hn} - Z_{ws}$ of the out-of-sample analysis under continuity of care) in all of the analyzed districts. This leads to $n_p = 10$.

These paths are found to be consistent with both the real execution of visits collected from the historical data of the analyzed provider and the outcomes of the patient stochastic model, which are used for the planning scenarios in the out-of-sample analysis under continuity of care (Table III). In district NPB, a particularly high difference is present between the paths and both the historical data and the scenarios for planning. The high utilization rates of nurses in this district (see *real* utilizations in Table IV) was probably the cause of a reduction in the number of visits actually provided by the provider with respect to the standard, particularly for long stay patients, to compensate overloading of nurses.

B. Assignments Under Continuity of Care

Very balanced workloads are obtained during the planning phase in each district. Indeed, the EV approach with the expected demands from the patient stochastic model (used for the out-of-sample analysis) gives $Z_{planned}$ equal to 0.0273 in district NPA, 0.0878 in district PA, 0.0707 in district NPB, and 0.0340 in district NPC. These values differ from a perfectly balanced workload $(Z_{planned} = 0)$ because a relevant part of the workload depends on previous assigned patients, whose demands evolve along with the time.

With regard to the execution, $mEVPI$ and $mVSS$ values lower than 0.05 are obtained in the first in-sample analysis (Table V). A negative $mVSS$ is present in district NPB; this holds because range Z is not the objective function of the mathematical programming model, even if strictly connected, and has to be intended as a null $mVSS$.

In the second out-of-sample analysis, very higher Z_{eev} values with respect to Z_{ws} are found in all of the districts (Table VI), underlining that the assignment problem is highly affected by the variability of patients' demands. A particularly high $Z_{eev} - Z_{ws}$ is obtained in district NPB, where a relevant difference between the paths and the estimates of the patient stochastic model used for planning scenarios is observed (Table III).

The HN analysis allows to divide the value of perfect information $(Z_{hn} - Z_{ws})$ from the non robustness of only considering the expected values $(Z_{eev} - Z_{hn})$. The HN approach has been executed with a number of scenarios n_s equal to 10 and 100. This last experiment $(n_s = 100)$ is highly time consuming for a practical real-time application in HC service providers (about 50–60 min for each week with about 20–30 new patients to assign, on a computer equipped with processor Intel Core i7 1.73 GHz and 6 GB of installed RAM). Moreover, with $n_s = 100$, the number of scenarios for the initialization at week 0 is reduced at 15, due to the significantly higher computational time when all of patients have to be assigned.

TABLE IV
MEAN UTILIZATION RATES

District	Number of nurses		Planned EV	Executed EV (real historical demands)	WS (real historical demands)	Real
NPA	8	Average	0.7963	0.8928	0.8954	0.8878
		Minimum	0.7854	0.7112	0.8600	0.7660
		25% quartile	0.7877	0.8649	0.8678	0.8254
		Median	0.7942	0.8798	0.8968	0.8752
		75% quartile	0.8009	0.9497	0.9194	0.9146
		Maximum	0.8126	1.0160	0.9371	1.1240
PA	3	Average	0.6370	0.6413	0.6402	0.9513
		1st	0.6005	0.6520	0.6420	1.0740
		2nd	0.6883	0.6253	0.6547	0.9040
		3rd	0.6222	0.6467	0.6240	0.8760
NPB	4	Average	1.0781	1.1368	1.1275	0.8591
		1st	1.1282	1.1693	1.2200	0.9067
		2nd	1.0657	1.3074	1.1154	0.7977
		3rd	1.0609	1.1336	1.0896	0.9024
		4th	1.0576	0.9368	1.0848	0.8296
NPC	5	Average	0.9077	0.9217	0.9342	0.9596
		Minimum	0.8878	0.8576	0.8672	0.7337
		25% quartile	0.9020	0.8680	0.9000	0.8344
		Median	0.9085	0.8960	0.9310	0.8580
		75% quartile	0.9183	0.9509	0.9566	0.9520
		Maximum	0.9218	1.0360	1.0160	1.4200

Mean utilization rates $\bar{u}_j$ under continuity of care in districts with more than one nurse: planned values of the out-of-sample EV approach and related execution in the real historical demands; values of the WS approach in the real historical demands (clairvoyant case); values with the real assignments implemented by the provider. Minimum value, maximum value, median and quartiles are reported for districts with more than four nurses.

TABLE V
IN-SAMPLE ANALYSIS UNDER CONTINUITY OF CARE

District	Z_{eev}	Z_{ws}	Z_{hn}	*mEVPI*	*mVSS*
NPA	0.1796	0.0974 (± 0.0132)	0.1360	0.0386	0.0436
PA	0.1677	0.1004 (± 0.0131)	0.1269	0.0265	0.0408
NPB	0.1129	0.1028 (± 0.0126)	0.1304	0.0276	-0.0175
NPC	0.1365	0.0778 (± 0.0145)	0.0988	0.0210	0.0377

Ranges Z, and *mEVPI* and *mVSS* values in districts with more than one nurse, for the in-sample analysis under continuity of care. The confidence intervals of Z_{ws} among the sample paths, used for choosing n_p, are also reported.

TABLE VI
OUT-OF-SAMPLE ANALYSIS UNDER CONTINUITY OF CARE EXECUTED IN THE SIMULATED PATHS

District	Z_{eev}	Z_{ws}	n_s	Z_{hn}	Z_{hn}- Z_{ws}	Z_{eev}- Z_{hn}
NPA	0.3197	0.0974	10	0.2813 (± 0.0418)	0.1839	0.0384
			100	0.3023 (± 0.0289)	0.2049	0.0174
PA	0.3213	0.1004	10	0.2988 (± 0.0787)	0.1984	0.0225
			100	0.3015 (± 0.0806)	0.2011	0.0198
NPB	0.5480	0.1028	10	0.2976 (± 0.0611)	0.1948	0.2504
			100	0.2615 (± 0.0339)	0.1587	0.2865
NPC	0.2296	0.0778	10	0.3604 (± 0.0871)	0.2826	-0.1308
			100	0.2840 (± 0.0528)	0.2062	-0.0544

Ranges Z and differences Z_{hn}-Z_{ws} and Z_{eev}-Z_{hn} in districts with more than one nurse for the out-of-sample analysis under continuity of care executed in the set of simulated paths (Z_{hn} is in terms of average value and half-width 95% confidence interval in brackets among the 6 repeats, and the differences are computed with the average value).

With both $n_s = 10$ and $n_s = 100$, the planning experiment is repeated in order to obtain a half-width 95% confidence interval of Z_{hn} lower than 0.1 in all of the districts. This results in six repeats.

Similar results are obtained with $n_s = 10$ or $n_s = 100$, indicating that the solution does not improve in the presence of such this order of magnitude for n_s. However, increasing the order of magnitude for n_s to improve the HN solution is not a viable approach because of its computational burden. In detail, high values of $Z_{hn} - Z_{ws}$ and low values of $Z_{eev} - Z_{hn}$ are obtained in almost all the districts (Table VI). A relevant $Z_{eev} - Z_{hn}$ is obtained only in district NPB: this is related to the differences between the simulated paths and the scenarios used for planning, which give the highest Z_{eev} in this district

TABLE VII
OUT-OF-SAMPLE ANALYSIS UNDER CONTINUITY OF CARE EXECUTED IN THE HISTORICAL DEMAND

District	Z_{eev}	Z_{ws}	n_s	Z_{hn}	Z_{hn} - Z_{ws}	Z_{eev} - Z_{hn}
NPA	0.3048	0.0771	10	0.3377 (± 0.0719)	0.2606	-0.0329
			100	0.2937 (± 0.0467)	0.2166	0.0111
PA	0.0267	0.0307	10	0.1689 (± 0.0637)	0.1382	-0.1422
			100	0.1089 (± 0.0408)	0.0782	-0.0822
NPB	0.3706	0.1352	10	0.3219 (± 0.0962)	0.1867	0.0487
			100	0.2650 (± 0.0531)	0.1298	0.1056
NPC	0.1784	0.1488	10	0.3147 (± 0.0621)	0.1659	-0.1363
			100	0.2855 (± 0.0448)	0.1367	-0.1071

Ranges Z and differences Z_{hn}-Z_{ws} and Z_{eev}-Z_{hn} in districts with more than one nurse for the out-of-sample analysis under continuity of care executed in the real historical demand of patients (Z_{hn} is in terms of average value and half-width 95% confidence interval in brackets among the 6 repeats, and the differences are computed with the average value).

TABLE VIII
PARTIAL CONTINUITY OF CARE AND REAL SOLUTION

District	Continuity of care (EV case)		Multiple operator assignment				No continuity between weeks		Z_{real}
			n_i=N_i=2		n_i=N_i=3				
	$Z_{planned}$	$Z_{executed}$ paths /historical/	$Z_{planned}$	$Z_{executed}$ paths /historical/	$Z_{planned}$	$Z_{executed}$ paths /historical/	$Z_{planned}$	$Z_{executed}$ historical	
NPA	0.0273	0.3197 /0.3048/	0.0044	0.2969 /0.2595/	0.0000	0.1389 /0.1532/	0.0016	0.1344	0.3580
PA	0.0878	0. 3213 /0.0267/	0.0059	0.1414 /0.0711/	0.0000	0.0972 /0.0362/	0.0026	0.0700	0.1980
NPB	0.0707	0.5480 /0.3706/	0.0000	0.2218 /0.0648/	0.0000	0.1345 /0. 1909/	0.0050	0.1219	0.1090
NPC	0.0340	0.2296 /0.1784/	0.0086	0.2782 /0.1641/	0.0000	0.1115 /0.1130/	0.0025	0.0808	0.6863

Ranges Z under continuity of care (EV approach of the out-of-sample analysis with $Z_{executed}$=Z_{eev}) and with partial continuity of care (multiple operator assignment and no continuity between weeks); assignments are executed both in the set of simulated paths and with the real historical patient demands. Z_{real} for the real operator workloads taken from historical data.

(Table VI). Hence, if the estimation is fine, the HN approach does not improve the solution (very low $Z_{eev} - Z_{hn}$ in the other districts). On the contrary, in the presence of this deviation, the EV approach does not work appropriately and the HN approach allows a significant improvement of the solution, even in the presence of a small number of scenarios (similar $Z_{hn} - Z_{ws}$ with $n_s = 10$ or $n_s = 100$).

Very similar results are obtained with the execution in the real historical data of demands (Table VII).

In conclusion, the high value of $Z_{hn} - Z_{ws}$ suggests that, if we could manage the complexity of the HN problem, the HN assignment would improve significantly.

C. Comparison With the Real Historical Assignments

Ranges Z_{real} of the $real$ solution (real nurses' workloads taken from historical data) are reported in Table VIII (last column). These ranges are compared with the outcomes of the analyses provided under continuity of care, because the provider aims at respecting this constraint.

Results show that the assignment models perform better than the real provider's practice. In detail, Z values under continuity of care (Tables VI and VII) are considerably lower than or similar to the Z_{real} values (Table VIII) in districts NPA, PA, and NPC. This demonstrates the benefits that can be derived from adopting the decision models proposed in this paper, even if the value of information associated with the assignment problem is still high. In addition, the higher workload balancing obtained with the proposed assignments leads to decrease of the number of overloaded nurses (with utilization rate $\overline{u}_j > 1$), while the $real$ assignments of the provider overload nurses even as their district is underutilized (see districts NPA and NPC in Table IV).

The benefits are made more evident because the outcomes of the planning models implement a perfect continuity of care, while in the $real$ condition this constraint is sometime violated to compensate for overloaded nurses or disruptive events we do not consider in the analysis (e.g., unavailability of nurses). Moreover, although the provider works with independent districts, separation among the districts is not always complete in practice, and nurses are sometimes allocated outside their territorial competence to compensate workload imbalances among districts (this is the reason of the different average utilizations for the $real$ case in Table IV, particularly for district NPB). Indeed, the continuity of care indices measured over the analyzed period for the $real$ case are $\lambda_P = 0.76$ and $\lambda_V = 0.70$.

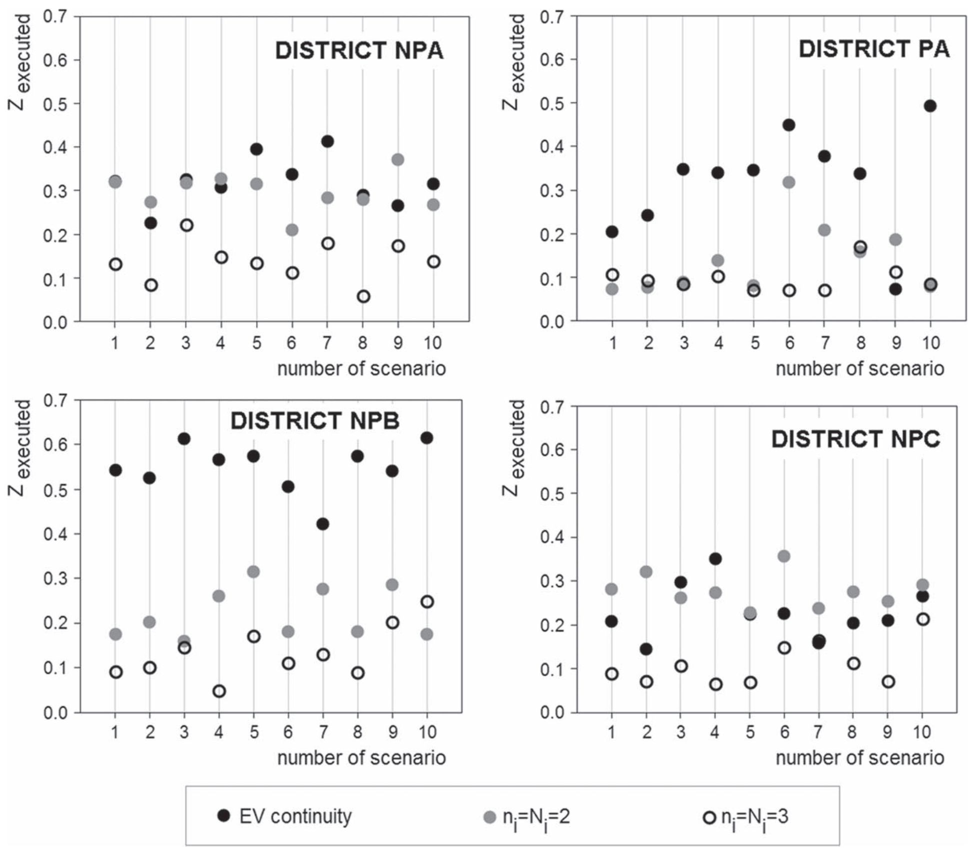

Fig. 2. $Z_{executed}$ for the EV approach under continuity of care (out-of-sample analysis executed in the simulated paths) and with multiple operator assignment.

Hence, even if the $executed$ and $real$ solutions should not be directly compared due to the disruptive events that the provider has to deal with in practice (holidays, illness, etc.), the implementation of the assignment models in the analyzed real case proved to increase workload balancing and continuity of care with respect to the usual practice of the provider.

D. Continuity of Care Relaxation

The analyses are performed planning the assignments with the expected demands from the patient stochastic model [25] and executing them in the n_p of simulated paths, similarly than in the out-of-sample analysis under continuity of care. Results of both the analyses with multi-operator assignment and without continuity between consecutive weeks are reported in Table VIII.

For more than one reference nurse, $Z_{planned}$ values decrease while increasing the number of reference nurses, assuming a null value for $N_i = 3$ in all of the districts. Also, $Z_{executed}$ values significantly decrease in the same direction. A detailed view of these results is reported in Fig. 2.

With regard to the absence of continuity between consecutive weeks, $Z_{executed}$ values are reported only for the real historical demands because of the different generations between the planning and the execution in the set of simulated scenarios. Indeed, the execution of the assignments without continuity between weeks is not possible in our experimental approach because the previously assignments are not maintained: if a patient is present in a week of a simulated scenario while he/she is not present in the planning phase, his/her reference nurse cannot be determined.

Very small $Z_{executed}$ values are obtained without continuity between consecutive weeks, meaning that the possibility of reassigning all of the patients at the beginning of each week provides more balanced workloads than the splitting of the demand (among two references nurses maintained over the weeks). However, this higher balancing is obtained while disrupting an important quality indicator of the provided service. The solution exhibits very low continuity of care indices: $\lambda_P = 0.50$ and $\lambda_V = 0.44$ for the planning; $\lambda_P = 0.55$ and $\lambda_V = 0.44$ for the execution. These values are considerably lower than those resulted from the $real$ assignments of the provider (where $\lambda_P = 0.76$ and $\lambda_V = 0.70$). In summary, if the continuity between consecutive weeks is neglected, the continuity of care indices are insufficient also when compared with the $real$ implemented assignments.

X. Conclusion

HC organizations require suitable tools to manage organizational activities, including the assignment of operators to patients, and to predict patients' demand evolution [5], [7]. Indeed, the possibility of estimating future patient demand [25] allows the implementation of decision-making models that assign patients to operators, while preserving the continuity of care and balancing the operators' workloads.

In this paper, we focus on these decision-making models. We propose a set of mathematical programming models to treat the assignment problem for different types of HC service providers, and we suggest the most appropriate model for different sets of provider characteristics. In particular, the continuity of care is considered in the proposed models, because this constraint represents the practice of many real HC service providers. Also, the variability of patients' demands is taken into account, with the goal of obtaining more efficient and robust plans, while in the literature this possibility is not exploited yet.

The proposed methodology proved to be effective when applied to a representative HC service provider. The numerical experiments demonstrated that the model produces assignments with perfect continuity of care and better workload balancing than those actually implemented by the provider, where human resources are assigned without the support of any decision model.

However, the analysis under continuity of care shows that a high value of perfect information is present in the assignment problem, due to high variability of the demands and the difficulties of capturing the real execution in the number of scenarios used for planning. As a consequence, a classical use of the stochastic programming technique is not able to significantly improve the balancing of nurses' workloads.

Therefore, other decision-making techniques under stochastic patient demands have to be investigated. Future work will be dedicated in completely using the available information (i.e., the entire probability density functions obtained from the patient stochastic model). First, a more accurate scenario generation will be investigated in order to exclude scenarios associated with a very low probability of occurring. Second, we are also working on both analytical structural policies and robust programming techniques which consider the entire density functions of the demand rather values extracted with a scenario generation [27].

Finally, we emphasize that the proposed models have been implemented in a software application currently in use by the analyzed HC service provider to assign patients to nurses. Feedbacks from the provider reveal an objective benefit in the planning, which derive from the assignments proposed by the application.

References

[1] S. Chahed, A. Matta, E. Sahin, and Y. Dallery, "Operations management related activities in home health care structures," in *Proc. INCOM Conf.*, Saint-Etienne, France, 2006, vol. 3, pp. 641–646.

[2] A. Matta, S. Chahed, E. Sahin, and Y. Dallery, "Modeling home care organizations from an operations management prospective," *Flex. Serv. Manuf. J*. doi:10.1007/S10696-012-9157-0.

[3] J. L. Haggerty, R. J. Reid, G. K. Freeman, B. H. Starfield, C. E. Adair, and R. Mc Kendry, "Continuity of care: A multidisciplinary review," *BMJ*, vol. 327, no. 7425, pp. 1219–1221, Nov. 2003.

[4] C. A. Woodward, J. Abelson, S. Tedfort, and B. Hutchison, "What is important to continuity in home care? Perspectives of key stakeholders," *Soc. Sci. Med.*, vol. 58, no. 1, pp. 177–192, Jan. 2004.

[5] V. Borsani, A. Matta, G. Beschi, and F. Sommaruga, "A home care scheduling model for human resources," in *Proc. ICSSSM*, Troyes, France, 2006, pp. 449–454.

[6] C. L. Cordes and T. W. Dougherty, "A review and an integration of research on job burnout," *Acad. Manag. Rev.*, vol. 18, no. 4, pp. 621–656, Oct. 1993.

[7] C. Castelnovo, A. Matta, T. Tolio, L. Saita, and F. De Conno, "Reforming Health Systems: Analysis and Evidence—Strategic Issues in Health Care Management," in *Proc. of 6th Int. Conf. Strategic Issues in Healtchare Management*, M. Tavakoli and H. T. O. Davies, Eds. Scotland: Univ. St. Andrews, 2004, ch. 11, pp. 135–151.

[8] V. De Angelis, "Planning home assistance for AIDS patients in the city of Rome, Italy," *Interfaces*, vol. 28, no. 3, pp. 75–83, May/Jun. 1998.

[9] C. R. Busby and M. W. Carter, "A decision tool for negotiating home care funding levels in Ontario," *Home Health Care Serv. Q.*, vol. 25, no. 3/4, pp. 91–106, 2006.

[10] M. Blais, S. D. Lapierre, and G. Lapierre, "Solving a home-care districting problem in an urban setting," *J. Oper. Res. Soc.*, vol. 54, no. 11, pp. 1141–1147, 2003.

[11] E. Benzarti, E. Sahin, and Y. Dallery, "Districting problem in home health care settings," *Decis. Support Syst.*, to be published.

[12] N. Lahrichi, S. D. Lapierre, A. Hertz, A. Talib, and L. Bouvier, "Analysis of a territorial approach to the delivery of nursing home care services based on historical data," *J. Med. Syst.*, vol. 30, no. 4, pp. 283–291, 2006.

[13] D. Boldy and N. Howell, "The geographical allocation of community care resources—A case study," *J. Oper. Res. Soc.*, vol. 31, pp. 123–129, 1980.

[14] A. Hertz and N. Lahrichi, "A patient assignment algorithm for home care services," *J. Oper. Res. Soc.*, vol. 60, no. 4, pp. 481–495, 2009.

[15] S. V. Begur, D. M. Miller, and J. R. Weaver, "An integrated spatial decision support system for scheduling and routing home health care nurses," *Interfaces*, vol. 27, no. 4, pp. 35–48, Jul./Aug. 1997.

[16] E. Cheng and J. L. Rich, "A home care routing and scheduling problem," Dept. Comput. Appl. Math., Rice Univ., Houston, TX, Tech. Rep. TR98-04, 1998.

[17] P. Eveborn, P. Flisberg, and M. Ronnqvist, "LAPS CARE—An operational system for staff planning of home care," *Eur. J. Oper. Res.*, vol. 171, no. 3, pp. 962–976, 2006.

[18] S. Bertels and T. Fahle, "A hybrid setup for a hybrid scenario: Combining heuristics for the home health care problem," *Comput. Oper. Res.*, vol. 33, no. 10, pp. 2866–2890, Oct. 2006.

[19] K. Thomsen, "Optimization on home care," M.S. thesis, Inf. Math. Model., Tech. Univ. Denmark, Kgs. Lyngby, Denmark, 2006.

[20] C. Akjiratikarl, P. Yenradee, and P. R. Drake, "PSO-based algorithm for home care worker scheduling in the UK," *Comput. Ind. Eng.*, vol. 53, no. 4, pp. 559–583, 2007.

[21] S. Chahed, E. Marcon, E. Sahin, D. Feillet, and Y. Dallery, "Exploring new operational research opportunities within the home care context: The chemotherapy at home," *Health Care Manag. Sci.*, vol. 12, no. 2, pp. 179–191, Jun. 2009.

[22] R. Ben Bachouch, M. Fakhfakh, A. Guinet, and S. Hajri-Gabouj, "Planification de la tournée des infirmiers dans une structure de soins à domicile," in *Proc. GISEH*, Lausanne, Switzerland, 2008.

[23] S. Martello and P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*. New York: Wiley, 1990.

[24] W. J. Hopp and M. L. Spearman, *Factory Physics*, 3rd ed. Homewood, IL: Irwin, 2007.

[25] E. Lanzarone, A. Matta, and G. Scaccabarozzi, "Workload estimation and balancing in home care organizations," *Prod. Plan. Control*, vol. 21, no. 1, pp. 3–25, 2010.

[26] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*. New York: Springer-Verlag, 1997.

[27] E. Lanzarone and A. Matta, "A cost assignment policy for home care patients," *Flex. Serv. Manuf. J.* [Online]. Available: http://www.springerlink.com/content/jg4m930413282262

Ettore Lanzarone received the Ph.D. degree in bioengineering from the Politecnico di Milano, Milan, Italy, in 2008, and the Master's degree in biomedical engineering *cum laude* from the Politecnico di Milano, Milan, in 2004.

He is a permanent Researcher at the Institute of Applied Mathematics and Information Technology of the Italian National Research Council, Milan, Italy. He is a Teaching Assistant of mathematical analysis at the Politecnico di Milano, Milan, Italy. In the healthcare area, his research interests include stochastic models for estimating the demand and planning the activities in healthcare structures, with particular reference to the home care services. Moreover, he also deals with scheduling algorithms for the manufacturing industry, with parameter estimation in dynamic systems and with the study of the cardiovascular fluid dynamics. He has published more than 20 papers in international journals and conference proceedings.

Andrea Matta received the Master's degree in industrial engineering from the Politecnico di Milano, Milan, Italy, in 1997.

He is an Associate Professor at the Department of Mechanical Engineering, Politecnico di Milano, Milan, Italy, where he teaches manufacturing, computer-aided manufacturing and manufacturing systems. He was a Visiting Professor at the Laboratory Productique et Logistique, Ecole Centrale Paris, Châtenay-Malabry, France, and at the Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA. He was a Professor of Manufacturing and Computer-Aided Manufacturing at Tongji University, Shanghai, China, in 2009.

His research interests include analysis, design, and management of production and service systems. He has published more than 80 papers in international journals and conference proceedings. He took part in several internationally funded projects, including: EMC2, MUPROD, VISIONAIRE, PROMISE, DRIVE, and MODFLEXPROD.

Evren Sahin received the Master's and Ph.D. degrees in industrial engineering from the Ecole Centrale Paris, Châtenay-Malabry, France, in 2001 and 2004, respectively.

She is an Associate Professor at the Laboratory of Industrial Engineering, Ecole Centrale Paris, Châtenay-Malabry, France, where she teaches operations management and supply chain management. She was a Visiting Researcher at Massachusetts Institute of Technology, Cambridge, MA, in 2008.

Her research interests include the design and analysis of production and distribution systems for products and services. She is interested in optimizing service processes, with a particular focus on health care. She has been involved in several national projects with hospitals, emergency medical services, and home care services.