The use of workload measurement as a total for nurse manpower planning

Citation for published version (APA):
Kusters, R. J. (1991). The use of workload measurement as a total for nurse manpower planning. In Operational research applied to health services (pp. 201-213). (Prace naukowe Instytutu Organizacji i Zarządzania Politechniki Wrocławskiej; Vol. 61), (Seria Konferencje; Vol. 20). Officina Wydawn. Polytechn..
THE USE OF WORKLOAD MEASUREMENT AS A TOTAL FOR NURSE MANPOWER PLANNING

In this paper the subject of allocating nursing staff to wards will be looked at in its context. Allocating nursing staff is one of the activities that are concerned with questions regarding demand and supply of resources in the hospital organization. It will be argued that these activities cannot be looked at separately, but have to be seen within a single unifying framework. In this paper such a framework will be outlined. Next the position of nurse staff allocation within this framework will be considered. From this position the need for a workload measurement instrument will be derived. Requirements for such an instrument will be identified and finally the use of such an instrument will be outlined.

1 Introduction

In this paper the subject of allocating nursing staff to wards will be looked at in its context. Allocating nursing staff is one of the activities that are concerned with questions regarding demand and supply of resources in the hospital organization. It will be argued that these activities cannot be looked at separately, but have to be seen within a single unifying framework. In this paper such a framework will be outlined. Next the position of nurse staff allocation within this framework will be considered. From this position the need for a workload measurement instrument will be derived. Requirements for such an instrument will be identified and finally the use of such an instrument will be outlined.

1 Eindhoven University of Technology, Department of Industrial Engineering and Management Science, Section of Management Information Systems and Automation, NETHERLANDS
2 Production control

As was mentioned before, the problem of allocating nursing staff has to be looked at within a wider framework. When we discuss the problem of allocating part of the available resources of an organization we have to realize that this problem cannot be treated isolated from the allocation of other resources on the one hand, and the demand that exists for these resources on the other hand. In short, we are talking here about the production control for the organization.

Production control busies itself with the coordination of all production activities [Bernard and Wijngaard, 1985]. In industry we normally associate production activities with the flow of materials through the organization, since the main activities of such organization is aimed at a physical transformation process. In a hospital however, the primary process is not aimed at the transformation of goods, but at the treatment of patients. Materials flow is of secondary importance. Not the flow of goods, but the flow of patients through the organization is the primary concern of hospital management.

Production control in a hospital can now be defined as containing those activities that are concerned with the coordination of all activities with regard to the flow of patients through the hospital. To put it in an other way, it is aimed at utilizing the (expensive) hospital resources fully on the one hand, and, on the other hand, tries to maximize the throughput of patients through the hospital.

These hospital resources are looked at as a combination of equipment, rooms and staffing (medical as well as nursing). Examples are operating rooms with equipment, nursing staff, anesthetic staff and medical staff and nursing wards with beds and nursing staff, the theme of this paper. Some notes will have to be made concerning this concept of production control in hospitals. First there is the idea that production control principles don’t with the personal care that is due to a patient. Patients are not goods and have to be treated accordingly. Therefore production control principles are not applicable to hospitals. This line of reasoning is a fallacy. The personal treatment that a patient has the right, quit rightly, to expect, has nothing to do whatsoever with the control that is exercised on the movement of patients as a group and on the allocation of resources. This kind of control takes place outside the view of the patient. The only effect patients notice is a better availability of the resources they require.

A second point that has to be made concerns the goals that production control aims at. It is easy to see that the two above mentioned goals, maximize patient throughput and maximize resource utilization, are
mutually incompatible. Take the example of emergency admissions. The hospital has to be able to admit urgent patients. However, since it is not known how many emergency admissions will occur, a tradeoff will have to be made. If many beds are reserved for emergency admissions, the chance of an emergency being refused through lack of capacity is low, but so is the average utilization of these emergency beds. This means, that in making decision regarding the allocation of resources and the throughput of patients a balance will have to be struck. The need for this balance is one of the main reasons for treating the subject of this paper, the allocation of nursing staff, within a production control framework.

A final note that has to be made regarding the occurrence of the word "coordination" in the definition of production control in hospitals. In this definition it is stated that production control is concerned with the coordination of all activities that have anything to do with the flow of patients through the hospital. That these activities have to be coordinated doesn’t mean that they have to be centralized. In fact, the hospital is such a large and complex organization, that to attempt such a centralized control would be nearly impossible. The control problem, considered as a whole, is simply too complex to allow such a centralized treatment. In one way or another we will have to arrive at a decomposition of this problem. This has to be done in such a way that the resulting, decomposed parts of the problem are small enough to allow an effective and efficient control while the central theme of coordination is yet achievable.

We will now take a closer look at this decomposition idea. First the general concept of decomposition of production control in a hospital organization will be examined. The resulting framework will be used to locate the main subject of this paper, nursing staff allocation. Such a decomposition can take place along several lines. The most obvious one is according to time. Production control decision can be classified according to the time scope over which they have effect. Here, decisions are called of a "higher level" when this time scope is longer. On each level a match between demand and supply of resources has to be achieved. In this regard Plossl and Welch (1979) make a distinction between the planning of the availability of capacity and the planning of the utilization of the capacity. It is clear that these two plantypes have to be in alignment to each other. Normally the following levels are distinguished:

- Long term planning.
  This planning level is aimed at determining the size of the facilities and the type of services to be rendered.
Middle term planning.
On this level a global assignment of resources to product types is made.

Short term planning.
On this level the actual production is planned, balancing available resources and actual demand.

This planning sequence is followed by execution of these plans on a shop floor level. Plans however have a tendency to fail, so at this level a shop floor control has to be inserted in order to provide for deviations from the plan. This gives us the configuration as is depicted in figure 1.

This decomposition sizes down the control problem into smaller parts that are easier to handle. Questions regarding "what has to be done", and "what is available to do it" are answered on each level in cohesion, while the level of detail increases as the time span over which the decisions are made shortens. The required coordination is achieved because each higher lever sets a framework within which the decision on the lower level have to be made.

Figure 1: decomposition of the control problem over time.

If we look at figure 1, we see that decisions regarding shop floor control are taken centrally. In reality is not very likely. A hospital is too large and complex an organization to allow this. This complexity is caused be the great number of parties involved in daily decision making and by the stochastic character of many of these decision. This means that further decomposition of the control problem is necessary. In some way one will have to distinguish units in the hospital organization form a production control point of view.
This decomposition has to be chosen in such a way that suboptimization per department is prevented. This suboptimization is caused by each department pursuing its own objectives regardless of the effect on the organization as a whole. In order to prevent this demarcation lines have to be chosen such that the resulting departments are capable of fulfilling their assignments independent from the other units. The effect of control measures taken in such a department must be local and may not influence other departments. [Bertrand and Wijngaard, 1985]. Furthermore, such a decomposition must not only result in a proper demarcation of functional competency, but also responsibility for the throughput of the patients has to be allocated. This idea, that the throughput of patients through the hospital is something to managed is often lacking in hospitals. All this makes it possible to achieve a balance between the supply of and the demand for resources. To achieve this it is necessary to be able to make realistic agreements with these departments regarding the execution of their assignments. [Bertrand and Wijngaard, 1985]

A department that fulfills these requirements is also called a production unit. The actual decomposition into production units to choose in a given situation can now be determined from these requirements. In an earlier paper [Kusters, 1987] I presented arguments for a decomposition where the outpatient departments, the outpatient operating rooms, the laboratories and the investigative departments were distinguished as separate units from a production control point of view. In this paper the clinic, combined with the inpatient OR was treated as a separate production unit. It is however possible that in larger
hospitals this large unit is divided into two or more smaller units (groups of nursing wards), thus losing some scale advantages while gaining on ease of control. We argued that the short term planning and the shop floor control in a hospital can't take place centralized because of the inherent complexity of the control problem at these levels. This argument doesn't hold for the higher levels. The long and middle term planning will be done centrally. This is not only possible, since the decision to be take at this level are not yet too complex to prohibit this. It is also necessary, since at this level no decomposition into relative autonomous units is possible. In figure 2 the resulting control configuration is depicted.

3 Nurse allocation in the production control framework

If we want to place nurse allocation in this production control framework, it is clear that we will only be depicting part of the picture. Nurse allocation, as one of methods by which the available resources of the hospital organization are directed, cannot be seen loose from the demand that is placed upon it. As was mentioned above, questions regarding supply of and demand for hospital facilities always have to be looked at in conjunction. The direct opposite to nurse manpower is the demand that is excersised on this capacity by inpatients. If we want to look at nurse manpower allocation and see the whole picture, this means that in each level the demand for this resource, and the control that is exercised on this demand, has to be taken into account as well. The result of this approach is shown in table 1. On the long term, we are talking here about a range of several years, the size of the hospital is determined together with the type of services that the organization is going to provide. From the demand side we get an overview of the type of demand that is going to be catered for. On the supply side this means for nurse manpower planning that the size of the nurse department, in full-time equivalent registered nurses and trainee, is determined. This gives us a frame of reference within which to take to further, lower level, decision. With a horizon of one year, production agreements give an indication of the workload that on average can be expected on the various nursing wards. Basically these agreements which are made between hospital management and the various specialisms deal with the total volume of activities during a year. On the basis of these agreements it is possible to make an allocation of nursing manpower to the wards. This
allocation is used every four to six weeks to make a duty roster, in which the duty shifts are divided over the available nurses, of course taking into account things as days of.

Table 1: stages in nurse allocation

<table>
<thead>
<tr>
<th>scope</th>
<th>control</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>org./unit</td>
</tr>
<tr>
<td>long term</td>
<td>hospital</td>
</tr>
<tr>
<td>&gt; 1 year</td>
<td>hospital</td>
</tr>
<tr>
<td>middle term</td>
<td>hospital</td>
</tr>
<tr>
<td>1 year</td>
<td>hospital</td>
</tr>
<tr>
<td>4-6 weeks</td>
<td>hospital</td>
</tr>
<tr>
<td>short term</td>
<td>unit</td>
</tr>
<tr>
<td>x weeks</td>
<td>unit</td>
</tr>
<tr>
<td>1-7 days</td>
<td>unit</td>
</tr>
<tr>
<td>1 day</td>
<td>unit</td>
</tr>
</tbody>
</table>

The available capacity as derived from the duty roster is one of the main input for admission planning. It presents information on the nursing capacity that would normally be available on a ward. It is only one out of several resources that have to be taken into account when planning the admission of patients into hospital. Others are available beds and available OR-capacity.

When planning patient admission it is tried to take the availability of nurse manpower into account, but since it is not the only restricting capacity, it is a rather flexible capacity and pressure on the organization towards a full utilization of the other resources is strong, it will quite often occur that a rescheduling of nurses (rescheduling-1 in the table) is necessary.

Finally when actual admissions (both planning and emergency) enter the hospital and the count of present nurse capacity is made (correcting for illness etc.) a final adjustment will have to be made. This shopfloor control will possible result in another rescheduling of nurses (rescheduling-2 in the table).
We have seen that it is quite possible to place the activities associated with nurse manpower allocation within this production control framework. Each decision on a higher level sets some degrees of freedom, thus narrowing down the scope of decision for the lower levels. The results of any level are used at a lower level, either by another manpower planning function or by a function associated with the control of the admittance of inpatients. The two sets of activities, the one aimed at assuring the supply of resources, the other aimed at assuring the demand for resources, are in this way intertwined so that a separation is not possible.

4 Workload measurement

We have seen that allocation of nurse and planning for the admission of inpatients cannot be looked at separately. However, this poses a problem. We see it is necessary to transfer results from one planning action to the other. It is for instance necessary to translate the results of admission planning into terms that have meaning when rescheduling nurses. This means that the need exists for a scale on which to measure both workload, as caused by inpatient, and available manpower, as presented by nurses. Furthermore, we need and instrument which makes it possible to translate both workload and manpower to this scale.

The most commonly used scale in this respect [de Wries, 1984] is working hours. It is easy to translate the available nursing manpower into the number of working hours that are made available. A certain requirement of working hours can easily be translated into the number and the type of shifts that are necessary to fulfill this demand. Translating the care requirements of inpatients into the number of working hours needed to fulfill this requirement is more a problem, but it is still quite possible.

If we look at measuring the workload as required by in-patients, we notice [de Vries, 1984] that a distinction can be made between:

a - an amount of work, depending on the type of patient,
b - an amount of work, depending on the number of patients, regardless of type,
c - an amount of work that has to be done anyway, regardless of patients.

The workload in the categories b and c can be determined, using traditional time studies. For category a some additional research has to be done. To determine this, an instrument has to be devised by means of which it possible to classify patients according to workload. Many such
Instruments exist. An example is [Murphy, 1978] which uses a scoring table as depicted in figure 3, in which four types of patients are distinguished with regard to the care needed:

I  minimal care or self care,
II  average care,
III  more than average care,
IV  continuous care.

On the basis of such a classification instrument it is possible to determine the workload in working hours caused by each of these types of patients on average. For this again use can be made of time study methods. An other option, copying the values from one other institution, is not advisable. Research in several Dutch hospitals showed [de Vries, 1984] that the workload per type patient varies between hospitals, and, even within hospitals, between wards.

<table>
<thead>
<tr>
<th>type of patient</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>independent</td>
<td>(  )</td>
<td>(  )</td>
<td>(  )</td>
<td>(  )</td>
</tr>
<tr>
<td>help bathing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>part. help posture</td>
<td>(  )</td>
<td>(  )</td>
<td>(  )</td>
<td>(  )</td>
</tr>
<tr>
<td>tot. help posture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>part. help feeding</td>
<td>(  )</td>
<td>(  )</td>
<td>(  )</td>
<td>(  )</td>
</tr>
<tr>
<td>infusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>observation 1-2 hour</td>
<td>(  )</td>
<td>(  )</td>
<td>(  )</td>
<td>(  )</td>
</tr>
<tr>
<td>constant observation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(weighing factors)</td>
<td>(v)</td>
<td>(v)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td></td>
<td></td>
<td>.5</td>
</tr>
</tbody>
</table>

Figure 3: patient classification instrument

One remark on the use of these time studies has to be made here. It appears [de Vries, 1984] that when the work pressure increase, non-patient related work tends to be postponed. When this is not taken into account when performing the time studies, a distorted picture might appear. This can be avoided by the addition use of a subjective evaluation during the time studies. Using only results from days which are evaluated as being "average" to determine the workload figures will then prevent the distortions from appearing.

Prerequisite for using these workload measurement instruments is their acceptance by the nursing staff. It they don't believe in the usefulness of such an instrument, it is very unlikely that the registration of data, needed to use it, will be done properly. After all, it are the nurses who have to classify all patients each day. If they don't trust the instrument it is more likely that data will be filled in later, or
not at all. In interviews with nursing staff four factors were mentioned that influence the acceptance of such an instrument.

- The level of detail should be correct. This means that the number of patient types that are distinguished should not be too high, nor too low. This is not an easy condition to fulfill, since different nurses will have different opinions on this subject.

- Classification must take place in a objective, yet simple way. This too poses problems. Classification should be simple, not time-consuming. This is understandable since work pressure on nursing wards is often quite high. The demand for objectivity on the other hand, mostly means that more questions are asked, costing more time for completion. For instance, one other method in use in the Netherlands takes more than an hour to complete each day.

- To each classification is attached an average workload figure. Nursing staff has to be convinced of the validity of these figures. This means that the time studies have to be performed carefully and with consultation of the nursing staff involved. Local for local determination of these figures will also increase the believe in the validity of these figures.

- Finally, even all these factors are taken into account, it will avail to nothing when no paper use is made of these data. This will be discussed further in the next section.

5 Using the results from workload measurement for staff allocation

For the remainder of this paper the assumption will be made that a workload measurement instrument is in use, so that data on workload can be gathered. Given this information it is now possible to look at manpower allocation separately. We will use table 1 as a starting point. When determining the size of the nursing staff, we can make use of data on average workload, caused by a certain type of service. These data give an indication for the amount of full time equivalent nurses that are needed given an identified package of services. In principle this should work. In the Netherlands however, such a mechanism is superseded by political arguments.

On the next lower level, the allocation of nurses over wards, we can make good use of these workload data. Given the average workload on a ward during the last period, it is possible to determine an allocation of nurses in such a way, that an on average equal workpressure is achieved on the ward. However, data on the previous period will not always suffice for this allocation. The new policies, as stated in production agreements, will have to be taken into account and translated
into their workload consequences as well. The historic data can here be used as a reference.

When the duty rosters are being made, no new information from the demand side is available, as can be seen in table 1. This means that the normal allocation can be used as a starting point here. One exception has to be made for the holiday periods. The effect of this period on the average workload can be seen from the historical data, thus giving a sufficient indication for the fabrication of the roster.

Up till now we only made use of average data on workload and season effects. These data didn’t take into account the actual situation on a given moment. When we come down to the short term planning and, finally, the shop floor control, it is not longer possible to ignore data on the actual situation. When we are reallocating nursing staff, in table 1 reallocation-1 and -2, this reallocation has to be based on a precise picture of the actual situation. In the following these two reallocation activities will be looked at closer.

Reallocation-1 (short term planning)
This reallocation activity tries, on the basis of the admission plan for the following day, to reallocate nursing staff so that an equal division of workload is obtained the next day. This means that in this activity it is necessary to look one day ahead. On day t an estimate for the workload on day t+1 has to be obtained. The workload on day t+1 is in principle caused by:

a - those patients that were in hospital on day t and will remain in hospital on day t+1,
b - those patients that are planned for admission on day t+1,
c - those patients that will be admitted with an emergency indication on day t+1.

Given the situation on day t, an admission plan for day t+1 and information on average numbers of emergency admission, it is now possible to predict the workload on day t+1. This prediction might be obtained subjectively from the nursing staff involved. It is shown [Warner, 1976], that such prediction are of a reasonable quality. However, these same experiments also show that it is difficult to obtain a continuing involvement from the nursing staff. Furthermore, such prediction will be used to reallocate nursing staff between wards. Reallocating on the basis of subjective estimates is likely to cause trouble. It is also for this reason that objectivity of workload measurement was required by the nursing staff (see above).

Given these considerations, a statistical model would be a better solution. That it is possible to design a model that predicts the workload on day t+1, starting from the situation on day t, with
sufficient accuracy was shown before [Kusters, 1988]. This model was based on the workload measurement instrument that was depicted in figure 3. It is based on a classification of patients into groups with a relatively homogeneous workload using preliminary diagnosis, age and sex. The workload attached to these groups was used for the prediction. If we look at the causes of workload on day t+1, it can be seen that this way the workload caused by the categories a and b can be predicted. If discharges for the next day are not known, meaning that category a is not known exactly, these can be predicted using the conditional distribution for length of stay given the time already spent in hospital. For category c we will have to make use of the average workload caused by an emergency admission.

Using this model it is possible to calculate expected value and variance of the workload on day t+1. This information can then be used to calculate the expected value of the workpressure, where workpressure equals workload divided by available manpower. By reallocating nursing staff it is now possible to even the workpressure over the wards.

Reallocation-2 (shop floor control)
The reallocation in the short term planning was partially based on predictions. These predictions have a tendency to be wrong. The predicted number of emergency admissions rarely follows expectations to the letter. Also, the number of nurses reporting sick is also hard to guess accurately. It therefore will often be necessary to perform another reallocation on the basis of the actual data. On day t+1 it is possible to measure the actual workload that is caused by the inpatients present in the wards. It is also exactly known how many nursing staff are available. Actual workpressure figures over the wards can be calculated. If necessary, a reallocation of nursing staff over the wards can follow, possibly using a special back-up teams of nurses who are used to irregular assignments. This second reallocation is not based on predictions, but on the actual situation.

One might wonder why, if this second reallocation has to be carried out anyway, is it necessary to carry out the first reallocation. The reason for this is, that it is very disquieting for nurses to have to move from ward to ward on a single day. It is better to announce such a move, when necessary, the day before. Furthermore, the main part of the workload is placed in the morning. When a large workload occurs without warning, it is rather late to do something about it. When the problem is recognized, and dealt with, so much time has passed that the reallocation will be late. All the more if one takes in account that a nurse who has been moved from one department to an other will need some time for adjustment.
It has to be noted, that this distinction between reallocation-1 and reallocation-2 means, that workload of patients has to be measured twice. Once, on day $t$, it is necessary to estimate the workload patients will cause the next day. The second time, on day $t+1$, the measurement of the workload of the patients in the wards will have to take place. This second measurement is used:
- to get an accurate picture of the present situation, needed for reallocation-1,
- to obtain the basic information needed for the long term planning,
- to obtain information for the calibration of the model,
- to get an impression of the accuracy of the predictions made on day $t$.
These last data can be used for feedback to the staff involved, in order to obtain a learning effect.

6 Conclusions

In this paper we looked at nurse manpower allocation. This theme was set in a larger production control framework. First this production control framework was presented and the nurse allocation problem was placed within this framework. From this it could be see that the allocation problem has to be looked at in conjunction with the admission planning of in-patients. As a link between these two planning sequences workload measurement was identified. Requirements on workload measurement instruments and the use of these instruments was discussed.

References


Kusters, R.J., Production control principles applied to the control of hospitals, Paper at the ORAHS meeting, Edinburgh, 1987.


W artykule omówiono zagadnienie rozmieszczania personelu pielęgniarskiego na poszczególnych oddziałach szpitalnych. Rozmieszczenie personelu pielęgniarskiego jest jedną z czynności ścisłe związanych z problematyką określania zapotrzebowania na usługi i ustalania wielkości zasobów w szpitalach. Dowiedzione zostanie, że czynności tych nie można traktować oddzielnie ale należy rozpatrywać je wewnątrz pojedynczej, ujednoliconej struktury. W artykule przedstawiono schemat takiej struktury oraz umiejscowiono w niej zagadnienie rozmieszczania personelu pielęgniarskiego. Następnie wykazano potrzebę zastosowania metody pomiarów obciążenia pracy, określono wymagania i przedstawiono zasady wykorzystania tej metody.