Wirtschaftswissenschaftliches Zentrum (WWZ) der Universität Basel

October 2011

Aggregation of Hospital Business Processes

WWZ Discussion Paper 2011/08
(D-123/D-127)

Markus Meier, Friedrich Rosenkranz

The Author(s):

Markus Meier, lic.rer.pol.
Forschungsstelle für Quantitative Planung (WWZ)
University of Basel
Peter Merian-Weg 6
CH - 4002 Basel

ma.meier@unibas.ch

Friedrich Rosenkranz, Prof.Dr.Dr.
Forschungsstelle für Quantitative Planung (WWZ)
University of Basel
Peter Merian-Weg 6
CH - 4002 Basel

friedrich.rosenkranz@unibas.ch

A publication of the Center of Business and Economics (WWZ), University of Basel.

Aggregation of Hospital Business Processes

Abstract

There are estimates that up to 30% of hospital costs are due to inefficiently coordinated hospital processes. As a result many hospitals have tried to model and to reengineer their business processes. These efforts have very often been abandoned, because the normally constructed total models of hospital processes could hardly cope with the rapid technological and medical progress as well as with changing staff. We discuss approaches for a qualitative and quantitative process modularization which improve the understanding of processes and enables better planned simulations. Various methods are discussed which allow a qualitative modularization on the basis of a disaggregated process graph. To cope with this modularization numerically simple semi-stochastic formulas are developed for the calculation of expected values and variances of cycle times and costs from micro-data up to the modular level. Thus a qualitative as well as quantitative discussion of hospital business processes on the modular level become possible.

Introduction

The costs for the health systems in the industrialized western countries are rapidly growing and have led to controversial political discussions and major efforts to control this development in a very problematic demographic environment. In 2008 the health care costs per capita were the highest in Norway, followed by countries like the US, France, Switzerland and Germany. In all these countries, more than 10% of GDP (US 16% of GDP) is used to finance the health system (OECD 2010). Hospital costs amount for more than 30% of health costs (Switzerland 48%) and have the fastest growing share. As a consequence, hospitals face a dilemma: on the one side the public expects rationalizations and cost control, on the other side patients ask for up to date medical treatment and care.

Different studies (Fisher et al. 2003; MBGH 2003; Milstein 2004) indicate that inefficient hospital processes again cause up to 30% of total hospital costs. Therefore, an analysis of the hospital value chains and efforts to reengineer processes seems to be a promising way to significantly reduce costs (Baldinger 2010, p.1-46). Considerable efforts to achieve such improvements have already been undertaken (Rieben et al. 2003; Walston et al. 2000, p. 1363-1388). However, most of these efforts failed to demonstrate convincing long-term benefits. The two main reasons for this result seem to be, first, the lack of standardization of hospital process models using so called reference models or building blocks and, second, the use of complete process models that led to a critical inflexibility of these models. Due to frequent technological and medical innovations hospital processes change quite often, be it gradually or even suddenly. An adaption of total models was often too cumbersome and too much dependent on individual staff. This was the case although today, at least in the German speaking countries, most hospitals work with SAP-systems (Muir et al. 2008) which constitute a certain standard in the area of process-data and -systems. They facilitate a migration to standardized process models or reference models. In the sequel we describe a modularization and standardisation approach which could form the basis for a more widespread use of process models. To demonstrate our approach we have selected three of the most frequent interventional/ surgical treatment areas. We focused on knee arthroscopy (International Classification of Diseases, 9th Revision, Clinical Modification ICD-9 CM: 80.27), on hip replacement (ICD-9 CM: 81.51- 81.53) and primary percutaneous coronary intervention (PCI) in acute ST-elevation myocardial infarction (ICD-9 CM: 36.0). These three treatment areas have been selected from approximately a thousand different hospital treatments because of their high volume and their high associated costs. The three chosen types of treatment constitute by number approximately 5% of the total hospital treatments in Switzerland (Bundesamt für Statistik 2009). Our process models have been developed in close cooperation with four hospitals situated in north-western Switzerland. The models developed are intended as reference models for these hospitals. Our process modelling and simulation work was also supported by IDS Scheer (ARIS) and COREL (Process 2009).

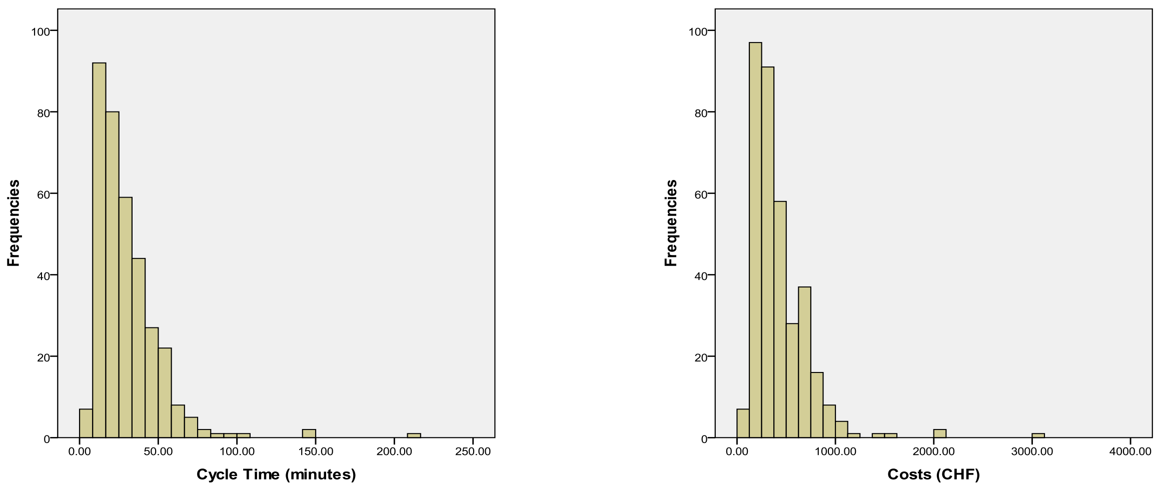


Fig 1. Duration and cost data for the process "arthroscopy".

Fig. 1 shows the duration (minutes) and costs (Swiss Francs, CHF) of the arthroscopy process at one of the hospitals mentioned based on 352 observed cases in 2008. Cycle time and costs include the data from incision to final suture without initiation and termination of anaesthetic procedures. The data shown were taken from the SAP-databases of the hospital. In other cases process data were estimated by experts (e.g. head physicians, scientists). To obtain such data, we followed a PERT methodology: optimistic, pessimistic and the values for the mode of individual activities were estimated and their expected values and variances were calculated assuming a beta-distribution (Neumann et al. 2002, p. 366-371). Most cycle times and costs directly observed have a distribution skewed to the right.

1. Modelling Standards and Structures

There exist some 10-20 different methods of graphical and logical process representation today. Graphical representations support the documentation and standardization of processes. Especially in the very initial design of a process, its description, visualization approaches facilitate the discussions of process properties and possible reengineering projects. Quantitative calculations and simulations supply detailed information for process improvements on the basis of these representations. A standard does not exist yet, but one would greatly improve the communication between organisations which, for example, cooperate in bench-marking studies by using a common standard. The presentation methods EPK and EPG (**E**vent oriented **P**rocess Chains or **G**raphs, Rosenkranz 2006) and BPMN (**B**usiness **P**rocess **M**odeling **N**otation, Havey 2005) have achieved a broader acceptance. For the sake of convenience, we use the EPG-notation in this paper. However, the same results would be obtained with other methods. Modern process software allows the flexible definition of graphical symbols and thereby the practical realization of different representation methods. On the computer process representations and data are kept in matrices and tables. They are also evaluated in this mode. Rows and columns of matrices correspond to nodes and/or arcs of a graphical representation. A graphical representation of a process shows how activities of a process depend on each other with respect to process transactions, logic, time and resources. In our case, a process transaction would be the trajectory that a patient is following on his treatment path. The representation and the embedded data give indications for expected but also for extreme process paths hospital-transactions may take. On this basis the graphical representations give indications for reengineering activities or structural process improvements, such as process accelerations, quality improvements or cost savings.

Fig. 2 shows the main symbols used by the EPG-notation (Gadatsch 2005, p.113-173; Rump 1999, p. 57-58), fig. 3 exhibits part of the graph which was established to describe the activities and events connected with a knee arthroscopy. A process starts and ends with events (rhombic starting node E_1, intermediate nodes E_2 and E_3, terminal nodes E_4 and E_5). Activities are represented by rectangular nodes with rounded corners like A_2 and A_3. They can be executed manually or in an automated fashion, e. g. using robots or by EDP-programs. They require time and resources and belong to organizational units (elliptic nodes Org. 1, Org. 2 and Org. 3). Their factor input (e. g. rectangular node for an information object or an amount of manpower or material etc.) leads to execution costs, their output to values and internal or external "sales". A sub-process A_1 is represented by the symbols of events and activities which are drawn on top of each other. A sub-process may contain a large number of individual events and activities. The circular node K represents a logic connector which may stand for a logical AND-condition (both A_2 and A_3 have to be executed in parallel), a logical XOR-condition (exclusively either A_2 or A_3 are executed) or, finally, the logically inclusive OR-condition (e. g. either both A_2 and A_3 or exclusively A_2 or A_3 are executed). The nodes are connected by arcs which either represent the control flow (directed arc), an organisational association (undirected arc) or factor input or output (directed arcs). The control flow determines the sequence of events, logical connectors, and activities which transactions encounter if they move through the process. In fig. 2 the points or "tokens" at the nodes of the graph represent transactions which are either waiting for service at an event or which are in process at the activities. Please note that a transaction which encounters an AND-connector is split into several daughter-transactions which move in parallel.

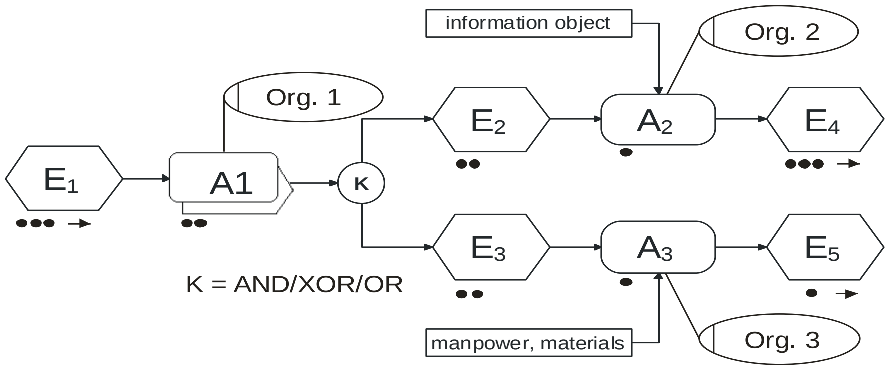


Fig. 2. Basic symbols of the EPG-notation.

Table 1 contains characteristic figures which give an impression of the complexity of the three processes described. Although fig. 3 shows only part of the arthroscopy process structure it is obviously rather difficult to understand the model logic from the detailed graph. Analyzing the structure of the same processes as it was ten years ago, we recorded the time, nature, location and costs of process changes which have actually taken place (backcasting). By asking our partners to estimate the time, location and costs of major process-changes to be expected within the next ten years we generated forecasts. Both, observed changes within the backcasting period and anticipated changes within the forecasting period did not only concern the medical content, but also the logistics and financial segments of the processes. Tab.1 shows the changes in cycle time and costs which have been or will be caused by the changes identified.

Table 1 Characteristics of process-structures

	Events	**Activities**	**Connectors**	**Arcs**	**Number of cases p.a. in Switzerland***
Knee arthroscopy	371	633	367	924	25’578
Hip replacement	971	1’733	967	2’424	16’575
Primary PCI	731	1’293	727	1’824	8’882

*Source: Bundesamt für Statistik, 2009

	Cycle Time [days] -10 years (change in cycle time %)	**Cycle Time** [days] present	**Cycle Time** [days][a)] +10 years (change in cycle time %)
Knee arthroscopy	3.8 (+23%)	3.1 (0%)	2 (-35%)
Hip replacement	15.8 (+34%)	11.8 (0%)	5 (-58%)
Primary PCI	5.1 (+21%)	4.2 (0%)	3.5 (-17%)

Source: Bundesamt für Statistik, 2009; [a)] estimated

	Area of change
Knee arthroscopy	For a long time, conventional radiography (X-ray) was the only imaging modality. The introduction of computer tomography (CT) and later magnetic resonance imaging (MRI) changed the hospital processes in the preoperative phase substantially and the performance of these techniques continues to improve. These newer imaging techniques allow imaging of soft tissue as well as appropriate diagnosis and careful planning of operational procedures.
Hip replacement	The development of minimally invasive operation techniques for hip replacement changed the intra- but also the post-operative care processes relevantly. This results in shorter procedure times and lower complication rates. This also resulted in reduced recovery time. Improved hip implants further reduced the recovery time; patients are able to become active and to lift weights soon after the operation.
Primary PCI	Percutaneous coronary interventions (PCI) in patients with acute myocardial infarction (MI) replaced non-invasive lysis therapy almost completely over the last few years. This is especially true for locations which are close to a heart catheterization laboratory. Ongoing advances in this field such as the invention of coronary stents, of drug-coated stents and of new anticoagulation medications (clopidogrel, prasugrel, bivalirudin, GpIIbIIIa antagonists, etc.) have further improved outcome and recovery of patients suffering from acute MI.

The data in tab. 1 demonstrate the complexity of hospital processes and are representative for many other medical treatments. Hospitals manage a great number of very complex processes. According to an estimate at the University Hospital Basel around 65’000 activities have to be planned and executed within more than a thousand processes/different treatments (Behrendt 2008, priv. comm.). It is extremely challenging to gain an overview over the processes and to plan their improvements on the micro-level, i.e., on the level of individual activities and events. We see process modularization and aggregation as a means to cope with this situation. Following a “Lego-strategy” it should be easier to develop and make available building blocks and reference models.

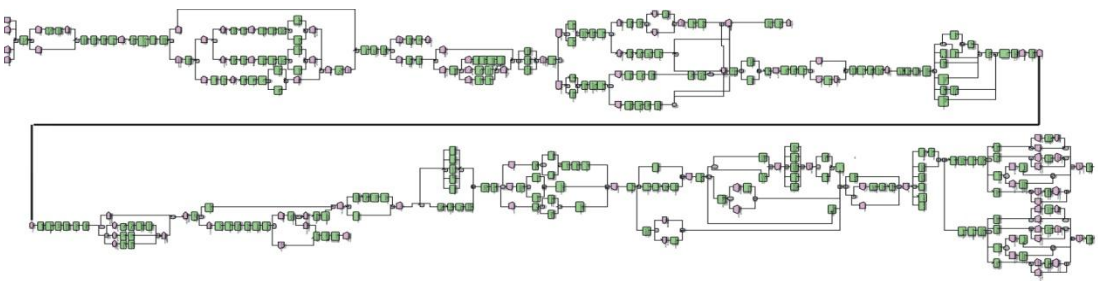


Fig. 3. Sub-process model of knee arthroscopy

The initial effort with this model building approach is higher than for the development of a total model ab initio. However, with the appropriate standards modules may be developed and applied at a number of hospitals making these efforts much more cost effective. In any case, the concept has to incorporate the possibility that hospitals attach individual process solutions to the system of reference models. If such modules are carefully defined and documented it should be much easier to change or update a module than a total process (Balzert 1998, p. 476). Network analysis as it is known from the scheduling area offers an approach to form sub-processes and modules (Jansen 2006). These networks are primarily based on graphs which only contain logical AND-conditions between the activities. In business processes also the XOR- and OR-conditions as well as feedback loops of activities and events have to be dealt with. Furthermore, scheduling-networks normally do not distinguish between paths in a graph which are sorted with respect to their importance or their probability of transactions taking place. However, this can be achieved using techniques similar to signal flow graphs, statistical path analysis or GERT (Elmaghraby 1977; Hair et al. 2006, p. 705-765; Mason 1953, p.1144-1156; Rosenkranz 1979, p. 185-236; Wright 1934, p. 161-215) which have to be adapted to the modelling of business processes.

Modularization consists of two types of analyses: First, the process structure is divided into sub-processes which are internally as well as economically and technically more homogeneous than the total process and can be handled independently (qualitative modularization). Second, detailed model data like cycle times, necessary resources or costs for activities have to be adapted or to be aggregated to the level of complete modules, where the qualitative as well as quantitative discussions take place (quantitative modularization). Quantitative modularization refers to quantitative model calculations that may be carried out on different aggregation levels.

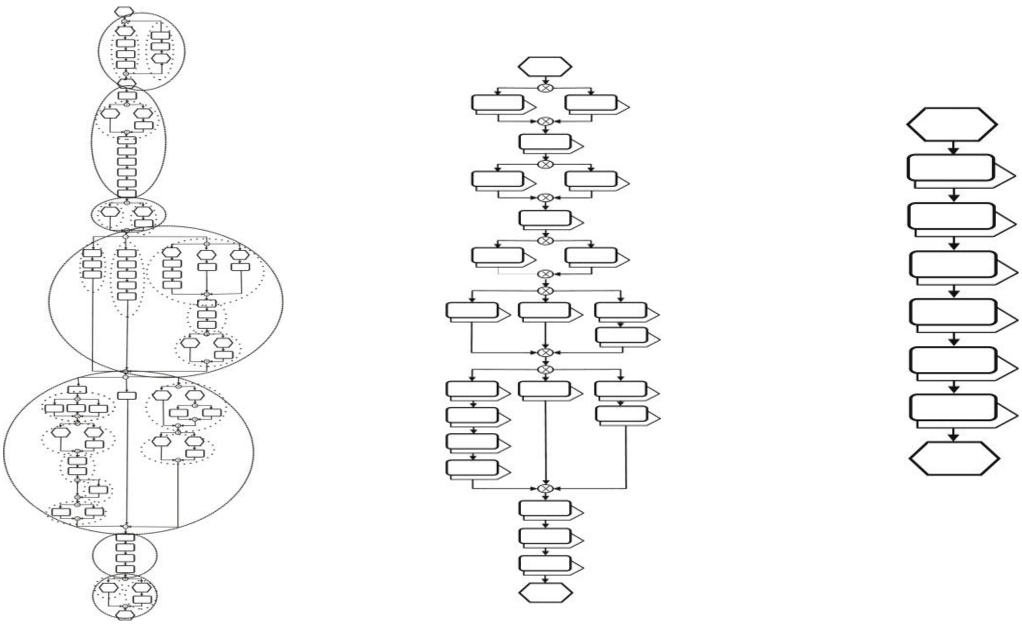
2. Qualitative Aggregation

A modularization of a hospital process along the boundaries of the existing organization structure is normally very simple (exist. Org. in tab. 2), because the observation of these boundaries is straightforward. Activities are under the control of already existing organizational units (fig 2). A disadvantage of a modularization along the existing organizational boundaries is the functional separation of doctors, nursing, purchasing, logistics and the economics departments. Such modularization does not support an integrated view of neither processes focused on patients or focused on different approaches to medical treatment. As a consequence the modules formed along existing boundaries lack transparency. They also frequently cause organizational problems with respect to the communication and coordination between the traditional departments.

Instead, the content of the modules may also be determined by cluster analysis (Cluster in tab. 2). Cluster algorithms try to minimize internal differences between the elements or nodes of a cluster and maximize external differences between different clusters. Normally these algorithms determine modules or clusters based on distances of a certain metric between elements or nodes of a corresponding graph. This concept has to be adapted to the process situation described. A distance of one was chosen for those nodes which are directly connected by an arc. The distance is assumed to be equal to two if the two nodes are separated by two arcs etc. If nodes are not connected at all, an infinite distance was assumed or for numerical calculations a distance of 100’000 was taken. These distances were assumed to be Euclidian. With a process-logic expressed by XOR- or OR-conditions, two nodes may directly be connected by one arc although they are not at all or only rarely attainable. In these cases the frequency, probability or importance of an arc was estimated and weighted with the distances between the nodes. For example if two nodes, say A_1 and E_2 in fig. 2, are connected by a XOR-condition K = XOR with a probability of 0.8 that E_2 follows A_1 and a probability of 0.2 that they do not follow each other, then we calculated as an “expected” distance the sum of 0.8 multiplied with a distance of one plus a distance of 0.2 multiplied with the defined infinite distance of 100’000. As a cluster algorithm we used the hierarchical Ward-algorithm which forms its clusters or modules using a least squares of distances approach (Backhaus et al. 2006, p.514-536; Ward 1963, 236-244).

The next approach tested was the well known Ford-Fulkerson-algorithm (FF in tab. 2) which is based on the max-flow, min-cut theorem (Cormen et al. 2007, p. 647-664). Again the underlying network-model is not directly compatible with the problem of process modularization and had to be adapted. In the Ford-Fulkerson picture process arcs have a limited and well defined upper flow-capacity. The standard algorithm is used to determine the maximum flow through a network. This flow has neither positive nor negative divergences, i.e. no flow units are lost or gained between the entrances and exits of a network. A minimum cut of the network defines a bottleneck which is equal to the maximum flow through the network. This cut determines the boundaries of a module. Fig. 4 shows an example with three hierarchical levels of modules. In order to adapt the standard algorithm to the problem of process modularization, capacities of one were assumed for all arcs which directly connect two nodes. The logical AND-condition causes a problem with this approach, because process transactions or flow units are split into daughter and grandchildren generation transactions (fig. 2). Thus they violate the flow continuity conditions in the nodes. This problem was solved by substituting the different paths between two AND-connectors by one arc of capacity one. Flows from or to OR- or

XOR-connectors were weighted with estimated path probabilities. This means that OR- or XOR-connectors split the incoming flow proportional to the probabilities that the following nodes succeed the connector. To illustrate this point let us assume two nodes, such as A_1 and E_2 in fig.2 which either have the connector K as a follower or predecessor. If the probability is 0.6 that E_2 follows A_1 then this corresponds to the flow from A_1 via K to E_2 and via A_2 to E_4. This results from the flow continuity conditions. With the same reasoning one obtains a flow of 0.4 from A_1 via K to E_3 and onwards until E_5 is reached. In both branches of the graph the unused capacity is 0.4 and 0.6, respectively. The minimal cut is in this case in the arc incident on K because it is saturated or has no additional capacity. In a next step all saturated arcs are removed from the process. The disjoint subgraphs thus created constitute the modules of level one. If the same algorithm is applied to these sub-graphs one obtains modules of the process which are disaggregated to the next lower level. Fig. 4 shows the results obtained for the modules on levels one to three.



level three level two level one

Fig. 4. Modules of the process "arthroscopy", Ford-Fulkerson algorithm.

Path betweenness-centrality (PBC in tab. 2) of an arc in a graph is a measure for the number of times a certain arc is contained in a path between two arbitrary nodes of a graph (Aier and Winter 2009, p. 150-163; Freeman 1979, p. 215-239). Betweenness-centrality is calculated for all arcs of a graph and sorted in descending order. Arcs which connect otherwise not connected sub-graphs have typically a very high centrality. If the arcs with the highest centrality are eliminated from the process graph one obtains sub-graphs or modules of level one. If the same procedure is applied to level one modules one obtains the modules of level two etc. The logical connectors contained in a process graph create again a difficulty compared to normal linear graphs: similar to the above described adaption of the Ford-Fulkerson algorithm, estimated path-probabilities were used for XOR- and OR-conditions to weight or diminish the connectedness of the corresponding arcs leaving the connectors. For AND-conditions all arcs leaving the connector are counted as different paths.

Just like syntax diagrams are used to describe the logic flow in computer programs and to identify possible subprograms, graph substitution (GS in tab. 2) is used to obtain the modules from given process graphs (Baresi et al. 2002, p. 402-429; Seol et al. 2006, p. 175-186). With this procedure again groups of arcs and nodes are substituted by modules. Whereas the methods described above work in a top down fashion, graph substitution works bottom up. It is carried out in several steps: in the first step all logical connectors are eliminated from the process graph. All groups of arcs and nodes which remain connected belong to a module of the lowest level. In the next step pairs of logical connectors (AND, OR, XOR) are introduced inside out again which allow the formation of a more aggregated module etc.

Tab. 2 summarizes the results obtained with the algorithms briefly described. It contains the three types of treatment, the number of major process changes observed within the backcasting and forecasting periods (i.e. 20 years in total), the number of modules resulting from the applications of certain method of aggregation and - in brackets – the number of module violations observed. A module violation occurs if a process change makes the adaption of more than one module necessary. The number of violations is defined as the number of modules which have to be adapted to the new situation minus one.

Table 2. Results of modularization analysis

[observed process changes due to technical and medical progress]	**exist. Org.**	**Cluster**	**FF**	**PBC**	**GS**
Knee Arthroscopy [7 changes]					
no. mod. level 1 (violations)	24 (12)	7 (1)	7 (1)	7 (1)	7 (1)
no. mod. level 2 (violations)	--	17 (3)	28 (2)	15 (3)	17 (1)
no. mod. level 3 (violations)	--	35 (5)	36 (2)	30 (4)	32 (2)
Hip Replacement [8 changes]					
no. mod. level 1 (violations)	25 (10)	8 (0)	8 (0)	8 (0)	8 (0)
no. mod. level 2 (violations)	--	19 (2)	30 (1)	17 (2)	19 (1)
no. mod. level 3 (violations)	--	38 (5)	39 (2)	31 (3)	35 (2)
Primary PCI (MI) [6 changes]					
no. mod. level 1 (violations)	27 (12)	7 (2)	7 (2)	7 (2)	7 (2)
no. mod. level 2 (violations)	--	21 (5)	32 (3)	19 (4)	20 (3)
no. mod. level 3 (violations)	--	42 (9)	42 (4)	38 (6)	36 (3)

The number of modules obtained for the currently existing organisation is self explaining. The number of modules on levels one, two and three was calculated with the mathematical algorithms briefly described. As one may see the existing organization corresponds best to the number of level two modules obtained with the mathematical algorithms. Major process changes within the backcasting or forecasting period are defined as changes at or above level three. In total 21 such changes were analyzed with respect to different level modularizations.

As one may see from tab. 2 the existing organisation requires a higher adaption effort than would be the case with the modules determined by mathematical algorithms. With respect to the number of violations that make module adaptions necessary graph substitution (GS) performed best followed by the Ford-Fulkerson algorithm (FF).

The methods described above may easily be implemented in existing process planning systems. However, due to political, personal and technical reasons (e. g. interfaces between different media, computer systems or existing organizational units) the results of the algorithms described should be viewed as a computer based proposal for a modularization. Staff involved in the running of a process model should be able to adjust the boundaries of the modules manually if necessary.

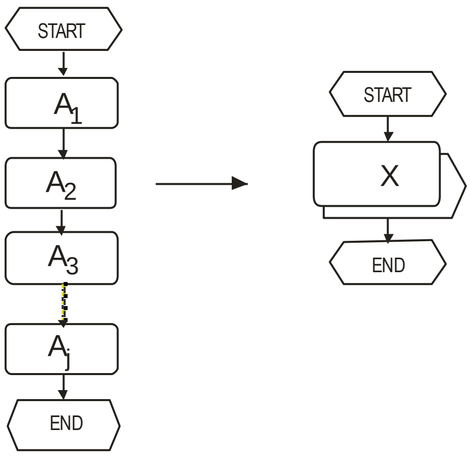
3. Quantitative Aggregation

A successful qualitative modularization may restrict expected process changes to isolated modules. On this aggregated basis it is easier to discuss possible structure improvements than on the micro-level. A user gets a better overview of the total process and does not get hung up in less relevant “process wall paper”.

In principle the same reasons speak in favour of an aggregated quantitative analysis. However, it is a priori not clear how data at the micro-level influence data on the module or aggregate level and vice versa. Quantitative analysis today is mostly carried out by either deterministic (“what if”) simulations or by stochastic simulations. In the latter case normally a Monte Carlo analysis is applied (Laguna and Marklund. 2005, p. 171-232). These approaches are pragmatic as well as flexible. Properly done, they provide unbiased solutions or different calculations of process performance even if the processes are large, stochastic, and of a complicated logical or functional structure. With modern software a stochastic simulation of, say, N = 50'000 different solutions and the calculation of the parameters of the random variables thus generated takes but a couple of seconds or minutes. However, structural discussions on the modular level and simulations on the micro-level normally do not properly fit together.

The following procedure aggregates micro-data for durations and costs of activities bottom up to the modular level. The aggregation described is semi-stochastic and non-parametric in the sense that it allows one to calculate expected values and variances without assuming certain statistical distributions. The calculations described below may substitute stochastic simulations on the micro-level and the user may see the causal influence of process changes more directly. In the case of deterministic calculations the expected values given below correspond to the deterministic values. Variances are not calculated in this case. One of the limitations of the approach should be mentioned beforehand: a hospital process may be viewed as a system of queuing lines of patients at the events and activities. Such stochastic waiting line problems are not described by the following calculations. The calculations assume that there are enough resources to service the transactions instantaneously. However, in the simulations described below this assumption was circumvented by using the existing capacities (e. g. personnel, equipment) of the processes. In case that capacities were altered the estimates of the duration of the activities were adapted correspondingly. In the sequel we designate by the symbols A_1, $A_2 \ldots A_j$ not only different activities, but also their duration, cost or value; $\varepsilon\{A_i\}$ is the expected value, $Var\{A_i\}$ its variance.

a) Aggregation of paths in series (fig. 5): a serial path of activities A_1, $A_2 \ldots A_j$ may be aggregated into one activity or a sub-process X by the calculation of the first two moments $\varepsilon\{A_i\}$ and $Var\{A_i\}$ via:



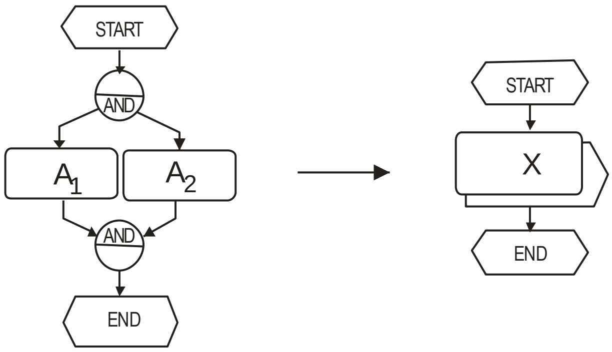
$$\varepsilon\{X\} = \sum_{i=1}^{j} \varepsilon\{A_i\}$$

$$Var\{X\} = \sum_{i=1}^{j} Var\{A_i\}$$

Fig. 5. Reduction of path in series

As said above: these results do not depend on the assumption of a certain distribution $F(A_i)$ or density function $f(A_i)$ of the activities. However, the variance formula assumes that all covariances or correlations between execution times or costs of different activities may be neglected.

b) Aggregation of AND-conditions: at a simple or nested AND-condition transactions are split up into next generation transactions. At the next AND-connector these transactions normally reunite, but the faster transactions have to wait for the slower ones. The aggregation of cycle time and the aggregation of costs or values have to be handled differently for AND-connectors. Cycle times are determined by the slowest transactions, whereas the costs always add up in all branches of an AND-condition. For two parallel paths we obtain fig. 6 and the following formulas for the cycle time, where j is used to specify different parallel path and k designates the longest path:



$$\varepsilon\{X\} \equiv \max_j\{\varepsilon(A_j)\} = \varepsilon\{A_k\}$$

$$Var\{X\} = Var\{A_k\}$$

Fig. 6. Reduction of AND-conditions (cycle time)

The above equations correspond to the formulas which are used for the determination of critical paths in stochastic scheduling networks PERT (Neumann and Morlock 2002, p. 376-371; Rosenkranz 2006, p. 157-164). Competing parallel paths after an AND-connector are substituted by the longest path. The variance of the path length of parallel time paths equals the variance of the longest path. Using these results one may calculate the probability that competing paths become critical or the longest by calculating $Prob\{(A_k - A_j) \leq 0\}$. To obtain numerical results one has to assume a certain type of density or distribution. As with PERT, the normal distribution seems to be the most convenient choice. The probabilities may thus easily be determined.

If this probability is not negligible, e. g. if

$$Prob\{(A_k - A_j) \leq 0\} \geq 0.2,$$

it follows that values for the expected duration and its variance may be corrected by calculating the distribution function (Hartung et al. 1995, 98-102):

$$F(X) \equiv Prob\left(\max_{i=[1,j]} A_i \leq X\right) = Prob(A_1 \leq X) \cdot Prob(A_2 \leq X), \dots, Prob(A_j)$$

$$F(X) = F_1(X) \cdot F_2(X) \cdot F_3(X), \dots, F_j(X)$$

$F(X)$ takes the form of a GEV-distribution (McNeil et al. 2005, p. 264-326). Unfortunately no general and simple formulas are available for this case. The results for the expected value and variance have to be calculated numerically by calculating the derivative of $F(X)$ and using $f(X)$ to determine the first two moments. We did not find it necessary in our experiments with the three hospital processes to correct the PERT-estimates given in fig. 6.

There is a difference in the reduction of costs or cycle times: since under AND-conditions all activities have to be executed, costs or value estimates are obtained by the following aggregation formulas:

$$\varepsilon\{X\} = \sum_{i=1}^{j} \varepsilon\{A_i\} \qquad Var\{X\} = \sum_{i=1}^{j} Var\{A_i\},$$

where A_i designate all activities belonging to an AND-condition.

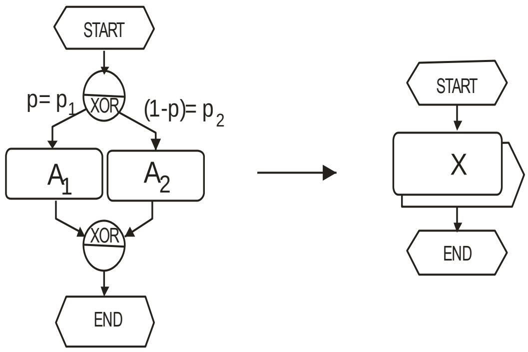
There is a major problem with the reduction of XOR- or OR-conditions in a process: very often the exits of the respective nodes depend on the evaluation of logical conditions. Example:

If DOCTOR IS AVAILABLE THEN DOCTOR PERFORMS A_1

ELSE ASSISTANT PERFORMS A_2

In this case one does not know how the logical condition should be evaluated numerically, since it depends on the availability of certain manpower resources. These conditions were evaluated under the assumption that it is possible to estimate frequencies or probabilities p_1, p_2...p_j with which the different cases occur.

c) Aggregation of XOR-conditions (fig. 7): we assume that neither the times or costs of different activities are correlated with each other, nor with the binomial random variable controlling the probabilities with which different paths are realized. Using the definition formulas for discrete random variables and elementary operations, we obtain the following results for the first two moments:



$$\varepsilon\{X\} = \sum_{i=1}^{j} p_i \varepsilon\{A_i\}$$

$$Var(X) = \sum_{i=1}^{j} p_i Var\{A_i\} + \sum_{\substack{i=1 \\ i\neq k}}^{j} \sum_{k=1}^{j} p_i p_k (\varepsilon\{A_i\} - \varepsilon\{A_k\})$$

Fig. 7. Reduction of XOR-conditions

d) Aggregation of OR-conditions: as is shown in fig. 8, OR-connectors are substituted by a combination of XOR- and AND-conditions. If the OR-connectors have several exits, probabilities for all possible combinations have to be estimated or to be assumed. In the case of only two activities like in fig. 8, we obtain three probabilities and branches which have iteratively to be reduced by rules b) and c).

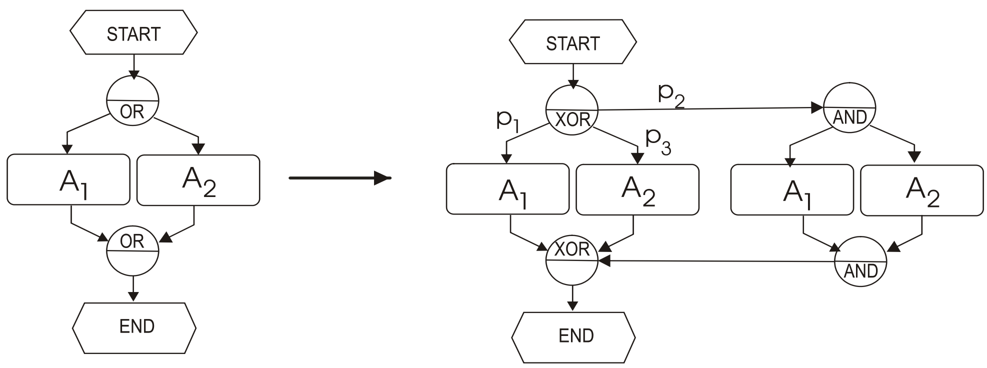
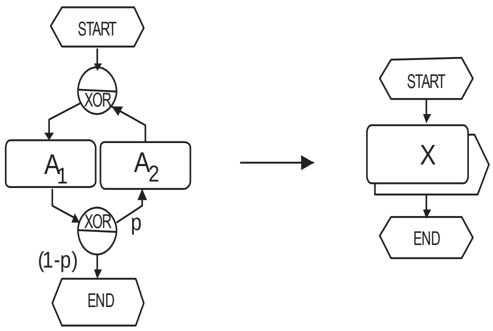


Fig. 8. Aggregation of OR-condition

e) Aggregation of cycles (fig. 9): The probability that a transaction leaves the cycle on its first turn equals $(1 - \mathrm{p})$. The probability that a second or third cycle occurs is equal to $(1 - \mathrm{p}) \cdot \mathrm{p}$ or $(1 - \mathrm{p}) \cdot \mathrm{p}^2$ respectively. Using the formula for the sum of a geometric series this results in the following reduction formulas:



$$\varepsilon\{X\} = \frac{(\varepsilon\{A_1\} + p \cdot \varepsilon\{A_2\})}{(1-p)}$$

$$Var\{X\} = \varepsilon(X^2) - \{\varepsilon(X)\}^2$$

$$Var\{X\} = (1-p) \cdot \sum_{i=0}^{n} p^i \{(i+1) \cdot \varepsilon\{A_1\} + i \cdot \varepsilon\{A_2\}\}^2 - \{\varepsilon(X)\}^2$$

Fig. 9. Reduction of feedback loops

A loop searching algorithm has to be applied to identify such loops in graphs automatically (Rosenkranz 1979, p. 185-236).

Tab. 3 shows a comparison of a Monte Carlo Analysis (N = 50’000) for the three processes based on empirical data of the cooperating hospitals and the solutions obtained with the formulas and procedure described above on the level of complete processes. Assuming a normal distribution of the result no significant differences were found on the 0.05-level for the estimators of expected value and variance. As described above the formulas uses existing capacities and depending waiting times. If capacities (e.g. personnel, equipment) will be reduced in the simulation models, the formulas will not agree with the monte-carlo results. The more the capacities will be reduced, the stronger the discrepancy. To compare the two methods under the new capacity restrictions, the depending waiting times and durations must be estimated again for the formula solution as done previously for normal activities.

Table 3. Comparison Monte Carlo with Analytical Formulas

Cycle time [days]	**Knee arthroscopy**	**Hip replacement**	**Primary PCI**
$\varepsilon\{X\}$ Monte Carlo	3.5	12.1	4.2
$\varepsilon\{X\}$ Formulas	3.5	12.1	4.2
$Var\ \{X\}$ Monte Carlo	0.2343	5.7069	7.0684
$Var\ \{X\}$ Formulas	0.2336	5.6803	7.1111
Costs [CHF]			
$\varepsilon\{X\}$ Monte Carlo	5’573.2	18’697.1	11’645.9
$\varepsilon\{X\}$ Formulas	5’573.1	18’697.0	11’642.0
$Var\ \{X\}$ Monte Carlo	153’527.10	1’240’677.19	2’909’623.04
$Var\ \{X\}$ Formulas	152’865.36	1’237’140.85	2’907’593.36

4. Disaggregations

The rules described above allow data aggregations to the level of process modules. If one discusses process improvements on the macro-level, the question arises immediately whether quantitative modelling may help to disaggregate these results down to the micro-level. Unfortunately this leads to a numerically underdetermined problem. The structure of the respective modules on the micro-level has to be specified and estimates of costs and times have to be supplied. The relations and reduction rules described above are still valid. However, they only constitute boundary conditions for many possible solutions. This means that only the consistency of the new micro-data with the modular data can be tested by aggregation.

5. Conclusions

Our experience with three frequently encountered hospital processes shows that such processes can be modularized using the appropriate algorithms (graph substitution). We were able to demonstrate that this modularization results in more robust and viable modules compared to a modularization along the existing organisational department boundaries. By applying the definition formulas for expected values and variances of five special cases (paths in series, AND-/XOR-/OR-conditions and feedback loops) we obtained reduction formulas by which data on a micro-level may be aggregated to a suitable modular level. These may facilitate aggregated discussions, be it with graphical structures or with data. Thus, we think that both, qualitative and quantitative aggregation methods should be implemented in advanced process planning software.

References

Aier S, Winter R (2009) Virtual Decoupling for IT/Business Alignment – Conceptual Foundations, Architecture Design and Implementation Example. BISE 2009 (2): 150-163

Backhaus K, Erichson B, Plinke W, Weiber R (2008) Multivariate Analysemethoden, 12th edn. Springer, Berlin

Baldinger M (2010) Anwendung der quantitativen Prozessmodellierung im Spitalwesen, Masters Thesis Basel, 45-46

Balzert H (1998) Lehrbuch der Softwaretechnik: Software-Management, Software-Qualitätssicherung, Unternehmensmodellierung. Spektrum, Heidelberg

Baresi L, Heckel R (2002) Tutorial Introduction to Graph Transformation: A Software Engineering Perspective. Springer, London

Behrendt H (2008) Process Consulting Basle University Hospital

Bundesamt für Statistik (2009) Medizinische Statistik der Krankenhäuser, Neuchâtel

Cormen TH, Leiserson CE, Rivest R, Stein C (2007) Algorithmen – Eine Einführung, 2nd edn. Oldenbourg, München

Elmaghraby SE (1977) Activity Networks, Project Planning and Control of Network Models. Wiley, NY

Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL (2003) The Implication of Regional Variation in Medical Spending: The Content, Quality and Accessibility of Care. In: Ann Intern Med 138(4): 273-287

Freeman LC (1979).Centrality in social networks, Conceptual clarification. Social Networks 1(3): 215-239.

Gadatsch A (2005) Grundkurs Geschäftsprozessmanagement. 4th edn. Vieweg Wiesbaden

Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL (2006) Multivariate Data Analysis. 6th edn. Prentice Hall, New Jersey

Hartung J, Elpelt B, Klösener KH (1995) Statistik, 10th edn. Oldenbourg, Munich

Havey M (2005) Essential Business Modeling, O'Reilly, Sebastopol

Jansen D (2006) Einführung in die Netzwerkanalyse. 3th.edn. VS Verlag, Wiesbaden.

(MBGH) Midwest Business Group on Health in collaboration with Juran Institute, Inc. and The Severin Group, Inc. (2003) Reducing the Costs of Poor-Quality Health Care Through Responsible Purchasing Leadership. 2nd edn. MBGH, Chicago.

Laguna M, Marklund J (2005) Business Process Modeling, Simulation and Design. Prentice Hall, New Jersey

Mason SJ (1953) Feedback Theory: Some Properties of Signal Flow Graphs. IRE 41(9): 1144-1156

McNeil AJ, Frey R, Embrechts P (2005) Quantitative Risk Management. Princeton, New Jersey

Milstein A (2004) Testimony to U.S. Senate Health, Education, Labor and Pension Committee. January 28.

Muir N, Kimbell I (2007) Discover SAP, 2nd edn. Galileo, Bonn

Neumann K., Morlock M (2002) Operations Research, 2nd.edn. Hanser, Munich

OECD (2010) [Organisation for Economic Cooperation and Development] OEDC Health Data, Paris

Rieben E, Mueller H, Holler T, Ruflin G (2003) Pfadkostenrechnung als Kostenträgerrechnung.

Ecomed, Landsberg

Rosenkranz F (1979) An Introduction to Corporate Modeling. Duke University Press, Durham.

Rosenkranz F (2006) Geschäftsprozesse: Modell- und computergestützte Planung, 2nd edn. Springer, Heidelberg

Rump FJ (1999) Geschäftsprozessmanagement auf der Basis ereignisgesteuerter Prozessketten. Teubner, Stuttgart

Seoul H, Kim C, Park Y (2007) Process Modularization: Concept and Algorithm. Concurrent Engineering, 15(2): 175-186

Walston SL, Burns LR, Kimberly JR (2000) Does Reengineering Really Work? An Examination of the Context and Outcomes of Hospital Reengineering Initiatives. Health Services Research, 34(6): 1363-1388

Ward JH (1963) Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58: 236-244

Wright S (1934) The Methods of Path Coefficients. Annals of Mathematical Statistics 5: 161-215