White Paper

QPR

QPR ProcessAnalyzer in Finnish Healthcare – Case Cardiology

Antti Manninen, QPR Software Plc.

Table of Contents

Introduction

In the early development of QPR ProcessAnalyzer, the product was tested using data from the cardiology treatment unit at Meilahti hospital in Finland. This white paper describes the test results from that study.

Cardiology is a field of medical science that studies and treats cardiovascular disorders, meaning disorders of heart and blood vessels. The nature of these disorders is naturally quite complex and most cases have to be individually treated, but there might still be some patient episodes that occur more frequently than the others and identifying these would help to streamline the treatment processes.

The cardiology data handled in this white paper consists of all visits to the first aid units at Meilahti and Maria hospitals, all visits to the Cardiology polyclinics at Meilahti and Maria hospitals, all visits to certain wards at Meilahti and all visits to Cardiovascular Research Unit (term CRU used from now on) in the years 2004 and 2005. The data is exactly the same as that collected by Nordic Healthcare Group (NHG) in 2006 in the so called Ihannesairaala (ideal hospital) project [1].

The purpose of this data selection is to identify longer treatment paths between these different units and to handle real life healthcare data. This enables validating the Automated Business Process Discovery (ABPD) algorithm in a realistic situation. Because the same data has already been used earlier by NHG in their study, the case is familiar with the consultants interviewed while conducting the study and the validation of the results becomes easier.

Research setting

Even though the data was selected to only consist of treatment activities in a few units in a limited time span of two years, the amount of data is still quite large. What causes even more problems to the clarity of the discovered process models is the heterogeneous nature of the data – there are multitude of totally different cases and trying to fit them all to one process model causes some confusion to say the least. This is why the data was filtered even further when validating the algorithm – the scope was selected based on some mutual factors that exist between many patient cases and the scoping could have also been done in a number of other ways.

For this thesis the cases were filtered so that only cases where the patients had at some point diagnosis codes I20–I25, I50 or R07 were selected. In order to reduce the effect of dispersion of the cases into many different process classes the filtering continued by selecting only cases with patient identifier of less than 2000. The diagnosis code based selection of cases was similar to that done in the NHG study; however that study did not limit the cases further. Once the patients having these special diagnosis codes had been found from the data, the specific patients were selected as the main data and the filter from diagnosis code was removed. This removes the problem that in many cases the diagnosis code changes during the process and simply taking the situation when the patient actually has a specific diagnosis code discards some of the earlier steps in the patient episode. These filters applied to the source data dropped the amount of cases from over 72000 to less than 350. The number of cases is still quite high after these limitations and the filtered data can be safely used as source data for the ABPD algorithm.

It is still likely that the resulting process model containing all the activities performed to the 350 identified cases becomes quite complex. The reason is that it contains every individual treatment event performed to every patient and some of the events or event combinations are quite rare. Special analyses are thus made by limiting the shown treatment events and flows between the events to the largest masses. In practice a Pareto rule is applied – after drawing the process model containing all the activities and flows, another process model is drawn by selecting only top 80% of the activities and flows. The result should follow the Pareto principle and consist of only a fraction of the original activities and flows.

Discovery results

The results presented in this chapter include the generic process model generated by the ABPD algorithm when showing all the activities and flows and the simplified process model generated when showing only the main activities and flows. Both of these models are analyzed and the applicability of QPR ProcessAnalyzer is evaluated in the case of the cardiology data.

The pictures below (Figure 1 and Figure 2) contain the beginning of the process model discovered from this data. The process model is quite complex and contains many flows between activities that occur only in a handful of cases. It is already clear at first glance that model like this is not very useful for thorough analysis, but it can be used to recognize the main flows between activities. Because of the high amount of flows between the activities, the labels indicating the amount of cases passing through each flow is not clear in this overall picture. However, the thickness of the arrow indicates the amount of cases. In this model the cases are divided quite equally between small individual activities and large flows only exist around the first aid polyclinics of Maria and Meilahti hospital, more specifically the units MRSPPKL and PÄ.

The picture here does not tell the whole truth of the generated process model. QPR ProcessAnalyzer can show flows related to an activity when user clicks it. This further information enables reviewing the flows and their volumes in higher detail. For a complex model like the one presented here this is one way of showing all the information. However, when simplifying the model also the flow labels indicating volumes become important.

An overall picture such as that presented by these figures is, however, good for making rough analyses of the whole situation. For example, when looking at Figure 1 it seems that the first aid polyclinics at both Maria and Meilahti hospitals are in a heavy use – 28% of the patients arrive to the first aid polyclinic at the Maria hospital and 27% to the respective polyclinic at the Meilahti hospital. An interesting remark is, however, that a relatively small amount of patients move from them to any other unit. The reason for this might be that the majority of patients are sent straight back home or that the patients are distributed evenly to the smaller units, thus making their flows less significant. The latter case can be evaluated by identifying more closely the amounts of patients moving to either of the cardiology polyclinics. As the pictures reveal, these amounts are not very high which indicates that majority of the patients actually just visit the first aid polyclinic and go home straight after this.

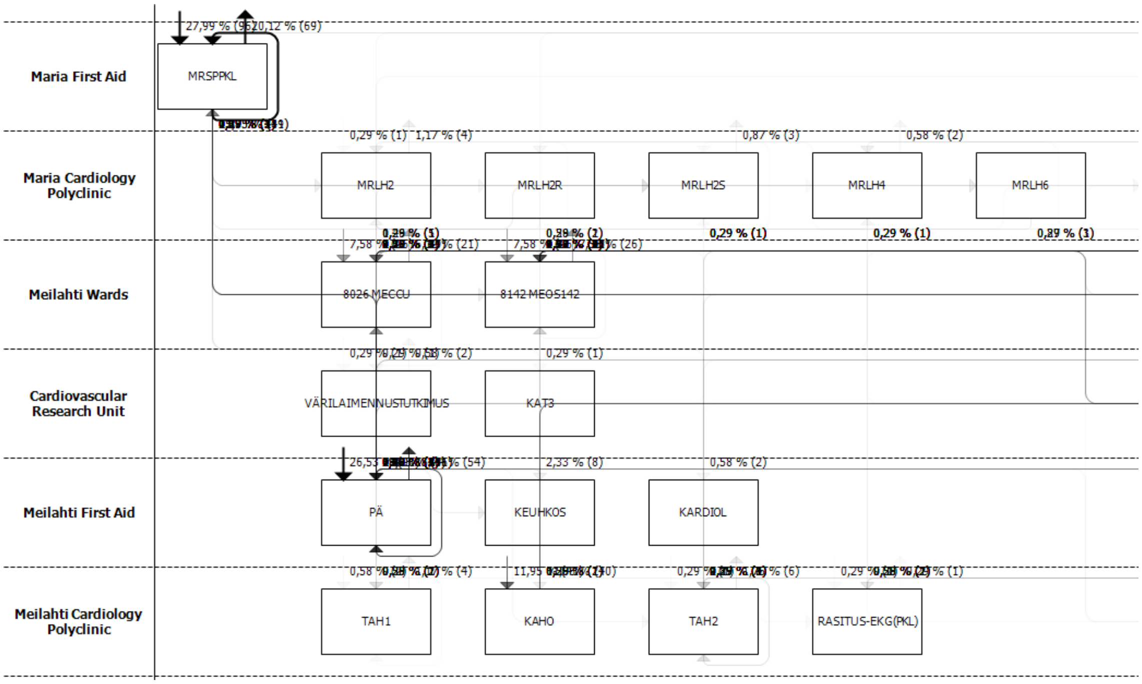

Figure 1: Cardiology data, discovered process model 1/2

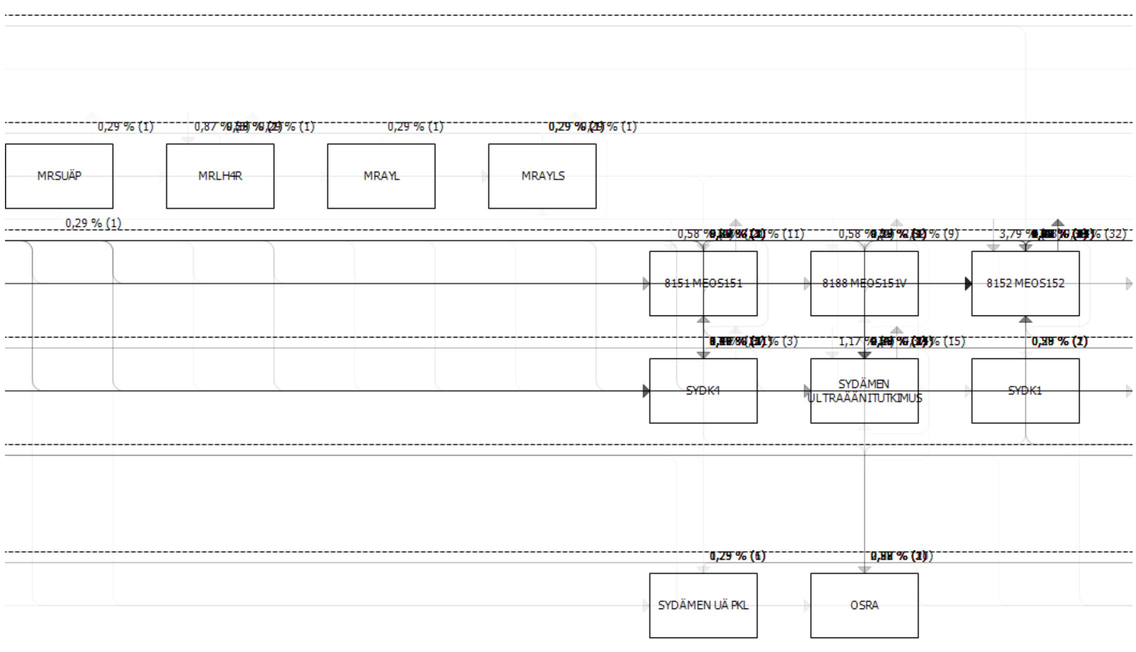

Figure 2: Cardiology data, discovered process model 2/2

The process model reveals also that none of the flows between activities are very large. The highest figures in these cases are around 25%, meaning that there is no single activity that all or even a majority of the cases go through.

When actually analyzing a process like this, the next step would probably be to separate the Meilahti and Maria first aid polyclinics and draw separate process models for both of them. However, as medical study is not the main focus of this white paper, I will not go into details here.

Pareto principle can still be applied here, even though there is no clear 80% majority in the model. High amounts in this case seem to be close to 10% of all the cases, so I will next present process model discovered from the same data but taking only account those flows where over 10% of the cases have been active. Figure 3Figure 1 presents this model.

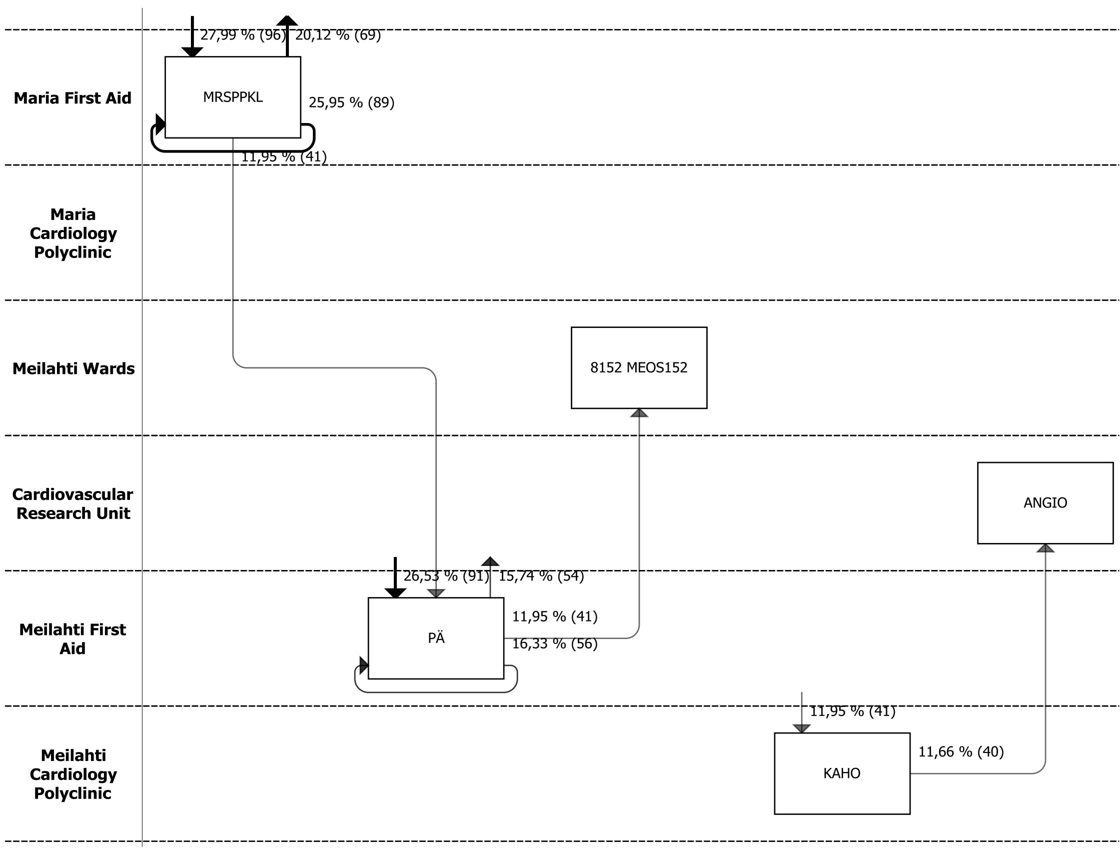

Figure 3: Cardiology data, mined process model with major flows

It is obvious that this process model is a lot clearer than that presented in earlier figures (Figure 1 and Figure 2). The major flows are naturally the same as those presented in the process model containing all cases, but this time the absence of the irrelevant small activities enables the focus of analysis to stay on the right track. Of course, if the point of the analysis is to examine a certain special activity, this will not be revealed in a Pareto analysis, but instead requires a special case constructed for this purpose.

To illustrate the capabilities of QPR ProcessAnalyzer further I will also present process model created from the same data, but this time the activities are simplified so that each organization unit contains only one similarly named activity. This reduces largely the complexity of process model presented in pictures earlier (Figure 1 and Figure 2) and should serve as demonstration of how the visual properties of process model can be largely affected by how activities are selected. The process model is presented in Figure 4.

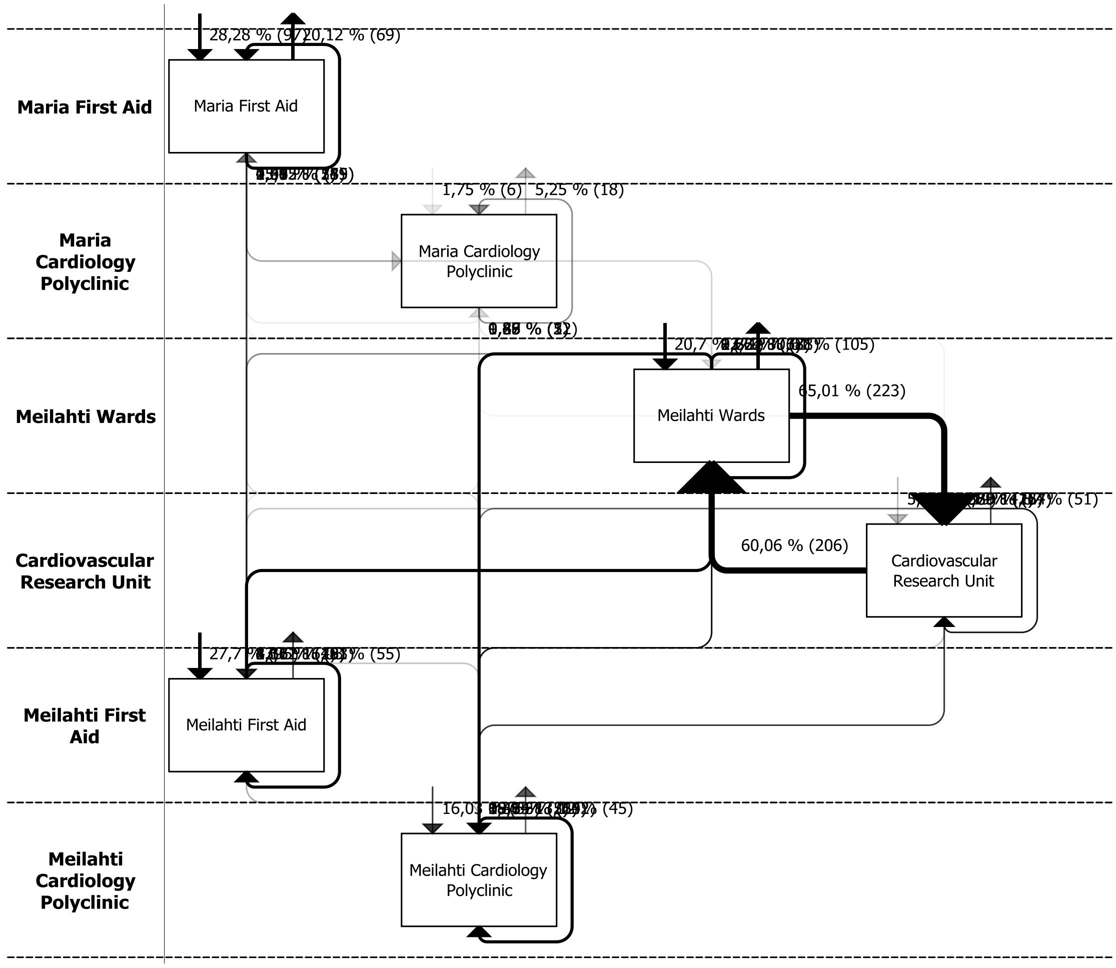

Figure 4: Cardiology data, simplified process model

When comparing this simplified process model with the process model presented earlier in this chapter, an interesting conclusion can be made. The major flows have moved from around the first aid polyclinics to the flows between the wards and the CRU. More specifically, the flows around the first aid polyclinics are of course still present as before, but combining the different wards and CRU's units together significantly increased their importance. Actually 65% of the cases found from the data move from ward to CRU and about 60% of the cases move back to wards. This highlights the importance of carefully selecting the data for QPR ProcessAnalyzer such that it can answer to the questions that are asked from it.

The nature of the data collected from the cardiology polyclinics is such that it does not contain very long treatment activity chains. The consequence of this is that large number of cases only consist of one activity – the patient arrives from somewhere outside the logs from which the data was taken and after inspection is sent to another unit. This can happen, for example, if the patient arrives from municipal healthcare. The problem with the data collected for this study is that it is not possible to identify whether the first activity for a case is actually the first activity the patient really encounters, or whether it is just the first activity that can be found from the analyzed log files. This causes some problems for interpreting the results generated by QPR ProcessAnalyzer.

Next Steps

To learn more about QPR Software's QPR ProcessAnalyzer please visit

http://www.qpr.com/Company/QPR-ProcessAnalyzer.htm

and contact:

Teemu Lehto, teemu.lehto@qpr.com, +358 40 546 0202

Antti Manninen, antti.manninen@qpr.com, +358 50 432 6231

References

1) Eklund F., Kämäräinen V., Tolkki O., Torkki P. (2007): Geneerinen prosessimalli ja integroitu kustannuslaskenta. HUS, Ihannesairaala project. In Finnish. Available online at http://www.hus.fi/default.asp?path=1,28,820,13120,17956,25572,25573,25574,25640

QPR Software Plc

QPR Software Plc is an international, highly regarded partner for enterprises and public sector in process development and business performance improvement. QPR's mission is to help people and organizations to take control of their business processes and achieve their goals.

QPR software has been implemented in more than 1,500 organizations across the globe and is provided in more than 20 languages. QPR was founded in 1991, has its headquarters in Helsinki, Finland and co-operates with an extensive network of talented partners in over 50 countries worldwide.

QPR Software Plc
Huopalahdentie 24
FI-003350 Helsinki, Finland
www.qpr.com

Tel. +358 290 001 150
Fax: +358 290 001 151

QPR Customer Care
Tel: +358 290 001 155
customercare@qpr.com

Quality. Processes. Results.