Multivariate Six Sigma: A Case Study in an Outpatient Pharmaceutical Care Unit.

Alba González-Cebrián[a]*, Marta Hermenegildo[b,c], Mónica Climente[b,c] and Alberto Ferrer[a]

[a]*Multivariate Statistical Engineering Group, Department of Applied Statistics, Operations Research and Quality, Universitat Politècnica de València, Valencia, Comunitat Valenciana, Spain;* [b]*Hospital Universitario Doctor Peset, Department of Pharmacy, Valencia, Comunitat Valenciana, Spain;* [c]*Facultat de Farmacia, Universitat de València, Burjassot, Comunitat Valenciana, Spain*

* algonceb@upv.es, hermenegildo_mar@gva.es, climente_mon@gva.es, aferrer@eio.upv.es

Multivariate Six Sigma: A Case Study in an Outpatient Pharmaceutical Care Unit.

Six Sigma strategies for process improvement are widely used in industry and manufacturing. The spreading tendency to gather process data about hospital activity is leading to an increase of process improvement projects in the healthcare context. The complexity of these databases requires upgrading the classical statistical Six Sigma toolkit. In this paper we present a Six Sigma project carried out in an Outpatient Pharmaceutical Care Unit at Hospital Universitario Doctor Peset in Valencia (Spain), where we illustrate the benefits of using latent variables-based models, specifically Partial Least Squares Regression (PLS), integrating them into the DMAIC phases of the project.

Keywords: Multivariate Six sigma; healthcare; process improvement; PLS; pharmaceutical care.

Introduction

The application and interest of process improvement in hospital environments has a growing tendency during last years [1]–[4]. Strategies such as Lean [5], Six Sigma (6S) [6] or their combination (Lean Six Sigma, L6S) [7], traditionally used in industrial or manufacturing sectors, are being widely used in other contexts, such as finance or healthcare. There is a bunch of existing work that already shows how 6S and L6S concepts can significantly improve process performance. In terms of hospital service, improving the performance can have multiple meanings: reducing prescription errors [4], reducing the waste of time [1], [8], [9], increasing patient's satisfaction [10], etc. Undoubtedly, a hospital is a complex environment with many parallel processes going on and affecting the very same issue. For instance, staff rotations, interdependencies between internal services and specific patient profiles, affect and define the optimal workflow that should be applied to each case and, therefore, data should reflect as accurately as possible, this reality. Thus, to improve the care of these patients using

statistical tools like the ones included in the Six Sigma toolkit, it becomes mandatory to deal with increasingly complex datasets. This issue becomes even more critical considering the tendency towards personalized medical care, where the patient becomes the focus of the caring process, which means that forthcoming process improvement should account for both information about patients and the hospital processes involving them.

This prospect of an increase in data availability and complexity along with a growing trend of customer-targeted services resembles the paradigm shift occurring in the manufacturing industry, known as Industry 4.0. The term Medicine 4.0 or Healthcare 4.0 is the literal extension of the Industry 4.0 concept towards the healthcare environment, pursuing a medicine guided by a preventive and personalized approach. With the 4.0 paradigm, it becomes an undeniable reality that the Six Sigma toolkit needs to be upgraded with machine learning (ML) tools and more sophisticated multivariate statistical techniques, such as latent variable-based models [11]. These tools can be really helpful to discover patterns, explore the data and obtain accurate predictions. A data-driven and efficient solution for these purposes is provided by Partial Least Squares (PLS) [12], [13] models. PLS is a latent variable-based multivariate statistical tool that has already been integrated into the Six Sigma toolkit in the industrial context [15]–[17], but not in the healthcare environment, as far as we are concerned.

This paper proposes the use of latent variable-based multivariate statistical techniques, such as PLS, into the Six Sigma statistical toolkit for healthcare processes improvement, illustrating their implementation into the DMAIC (Define, Measure, Analyze, Improve and Control) phases [18] of a 6S project carried out in an Outpatient Pharmaceutical Care Unit in the Department of Pharmacy at Hospital Universitario

Doctor Peset in Valencia (Spain). This unit provides prescription drugs and pharmaceutical care services to outpatients. The outcomes of the multivariate Six Sigma approach will be compared with the conclusions obtained by classical Six Sigma statistical tools, such as the ANalysis Of Variance (ANOVA). It is our belief that this approach encompasses the possibility of both, confirm usual suspects and discover new relationships that had not been considered, which is one of the added values of Six Sigma projects.

Methodology

This section is divided into two main parts. The first one will briefly introduce the 6S methodology. The second one will give some mathematical and theoretical background about the latent variable-based multivariate models, specifically, about the Partial Least Squares Regression (PLS) model [12], [13]. In order to illustrate the inclusion of PLS as a tool for Six Sigma projects, we followed a two-step procedure along this work:

1) To fit a PLS model that will point out interesting (both new and suspected) relationships between process inputs and outputs. The general overview of the complex relationships between X´s and Y´s provided by PLS weighting plot is a useful tool to drive the following steps in the Analyze phase. This provides a route map of what is worth studying in more depth.

2) To assess these potentially interesting relationships with traditional explorative tools.

Six Sigma DMAIC methodology

One of the key advantages of the Six Sigma projects is the use of the systematic approach provided by the DMAIC methodology [18] that splits the route map of any

project into five phases: define (D), measure (M), analyze (A), improve (I) and Control (C). According to authors in [17], each stage of the DMAIC procedure can be briefly defined as follows:

- Define: problem selection and benefit analysis.
- Measure: translation of the problem into a measurable form, and measurement of the current situation; refined definition of objectives.
- Analyze: identification of influence factors and causes that determine the critical to quality characteristics' (CQCs) behavior.
- Improve: design and implementation of adjustments to the process to improve the performance of the CQCs.
- Control: empirical verification of the project's results and adjustment of the process management and control system so that improvements are sustainable.

Partial Least Squares Regression

Latent variable models (LVMs) are multivariate statistical models specifically designed to analyze massive amounts of correlated data. The basic idea underlying them is that the real dimension of the studied process is smaller than the number of measured variables, which therefore appear correlated. With this assumption, new variables named latent variables (LVs) are created by combining the measured variables. These latent variables are very useful for the identification of driving forces acting on the system since they optimally describe the sources of data variability.

Partial Least Squares Regression (PLS) [12], [13] pursues the obtention of LVs that maximize the covariance explained between a $N \times K$ space of inputs ($\boldsymbol{X}$) and $N \times L$ outputs ($\boldsymbol{Y}$), where N is the number of observations, K the number of predictor variables

and L the number of output variables measured. Whereas Multiple Linear Regression (MLR) or other Machine Learning (ML) techniques focus on modelling the relationship between inputs ($\boldsymbol{X}$) and outputs ($\boldsymbol{Y}$), PLS provides not only a model for this relationship but also a model for $\boldsymbol{X}$, what provides unique properties. LVs computed in the PLS model represent the main driving forces linking the input to the output space. The PLS regression model structure is described by the following equations and illustrated in Supplementary Figure 1:

$$\boldsymbol{T} = \boldsymbol{X}\boldsymbol{W}^{*}$$

Equation 1

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{T} + \boldsymbol{E}$$

Equation 2

$$\boldsymbol{Y} = \boldsymbol{T}\boldsymbol{Q}^{T} + \boldsymbol{F}$$

Equation 3

Columns of $\boldsymbol{T}$ are the PLS score vectors, conforming to a matrix of dimensions $N \times A$, where A is the number of LVs of the model. These vectors explain most of the covariance of $\boldsymbol{X}$ and $\boldsymbol{T}$, and each one is estimated as a linear combination of the original variables with the corresponding weight vectors from $\boldsymbol{W}^{*}$ (Equation 1). Moreover, PLS does not only model the relation between $\boldsymbol{X}$ and $\boldsymbol{Y}$, but also models their projection onto the latent subspace of dimension equal to the number of LVs. This is the reason why PLS scores, $\boldsymbol{T}$, are simultaneously good summaries of $\boldsymbol{X}$ according to $\boldsymbol{P}$ (Equation 2) and good predictors of $\boldsymbol{Y}$ according to $\boldsymbol{Q}$ (Equation 3). Besides, the number of selected variables is related to the effect of the dimensionality reduction. The bigger the reduction, the fewer LVs (A), and the information not represented by these A LVs is stored in the error terms $\boldsymbol{E}$ (for inputs) and $\boldsymbol{F}$ (for outputs). Consequently, $\boldsymbol{E}$ and $\boldsymbol{F}$ become key indicators of the PLS model goodness of fit: the smaller the sum of squares of $\boldsymbol{F}$ is, the better the model is for the prediction, and the smaller the sums of squares of

$\boldsymbol{E}$ is, the better the model explains the $\boldsymbol{X}$-space. Usually, the number of latent variables is selected in such a way that $\boldsymbol{E}$ and $\boldsymbol{F}$ matrices can be considered to contain nothing but noise, keeping the meaningful information (signal) stored in the A PLS latent variables.

For a given observation, to evaluate the model performance projecting the n-th observation, $\boldsymbol{x}_n$, onto it, the Hotelling-T^2 in the latent space, $T^2_{\boldsymbol{x}_n}$, and the Squared Prediction Error (SPE), $SPE_{\boldsymbol{x}_n}$, are calculated:

$$\boldsymbol{\tau}_n = \boldsymbol{W}^{*T}\boldsymbol{x}_n$$

Equation 4

$$\boldsymbol{T}^2_n = \boldsymbol{\tau}^T_n \boldsymbol{\Lambda}^{-1}\boldsymbol{\tau}_{\mathrm{n}}$$

Equation 5

$$SPE_{\boldsymbol{x}_n} = (\boldsymbol{x}_n - \boldsymbol{P}\boldsymbol{\tau}_n)^T(\boldsymbol{x}_n - \boldsymbol{P}\boldsymbol{\tau}_n) = \boldsymbol{e}^T_n\boldsymbol{e}_n$$

Equation 6

where $\boldsymbol{e}_n$ is the residual vector associated with the n-th observation, $\boldsymbol{\Lambda}^{-1}$ the diagonal matrix containing the inverse of the A variances of the scores associated with the LVs, and $\boldsymbol{\tau}_n$ the vector of scores corresponding to the projection of the n-th observation $\boldsymbol{x}_n$ onto the latent subspace of the PLS model. T^2_n is the estimated squared Mahalanobis distance from the center of the latent subspace to the projection of the n-th observation onto their subspace. SPE_n gives a measure of how close (in a Euclidean way) such observation is from the A-dimensional latent space.

PLS model can be expressed as well as a function of the input variables by substituting Equation 1 into Equation 3:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{W}^{*}\boldsymbol{Q}^T + \boldsymbol{F} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{F}$$

Equation 7

where matrix $\boldsymbol{B}$ $(K \times L)$ contains the PLS regression coefficients stored by columns.

All PLS model parameters can be calculated sequentially using the NIPALS algorithm [12], which also handles missing data. This makes the PLS an attractive tool

for analyzing complex databases. Moreover, when the response variable is categorical, there is an adaptation of PLS that can be used for discriminant and classification purposes. This version is named PLS-Discriminant Analysis (PLS-DA) [19].

Results

The project timeline went from July 2018 to September 2019 (Figure 1). This section will follow the pathway defined by DMAIC steps, illustrating the results and the process of the project carried out.

[Figure 1 here]

Figure 1. Timeline of the Six Sigma project, indicating the data recording periods and implementation of changes.

Define Phase

The goal of this stage is to determine the improvement project potentially leading to a reduction of costs, an increase in the customers' satisfaction, etc. This implies a necessity of defining an observed problem to be tackled. Studying the process and its relation to the problem, an assessment of the costs and benefits of addressing the project goal can be evaluated. This provides a first clue about the necessities of resources, staff involved in the project and potential constraints. All these initial aspects were portrayed in the Project Charter (Supplementary Figure 2).

The focus of this Six Sigma project was related to the timing (waiting and attention times) of outpatients during their visit to the hospital´s Department of Pharmacy. According to the last outpatients' satisfaction questionnaire, performed between November 2016 and February 2017, half of the outpatients evaluated the waiting time as not short enough. This was reflecting an improvement opportunity based on the voice of the external customer, i.e., the outpatients. Besides, according to

the outpatient pharmaceutical care unit staff (nurses and pharmacists), the amount of work had been increasing during last years. Data about last year's (January – June 2018) agenda confirmed the voice of the internal customer (the staff), showing a clear and systematic overload of patients (Supplementary Figure 3). This overload was calculated daily, as the difference between attended and scheduled outpatients for each day.

Indeed, internal and external clients' voices were aligned towards the same direction: there was a consistent overflow of patients and hence long waiting times. Understandably though, changes on the organizational scheme would be carried out only under strong evidence supporting the need for improvement since other adjacent processes of the unit could be affected as well. However, the data recorded up until 2018 did not register the timing per patient, which made it very difficult to formally raise the patient complaints and redesign the unit's workflow. For this reason, with help of the pharmacist staff and considering both Voices of the Customers (VOC), a "Suppliers, Inputs, Process, Outputs, Customers" (SIPOC) diagram was outlined (Figure 2) to design a data collecting scheme.

[Figure 2 here]

Figure 2. SIPOC diagram of the workflow in the Outpatient Pharmaceutical Care Unit.

The project was led by a Six Sigma black belt with a high profile in the hospital pharmacy organizational scheme, and the technical team consisted of six members of the Outpatient Pharmaceutical Care Unit staff and two black belts with an engineering and statistical background. The Six Sigma project was championed by the chief of the hospital pharmacy department.

The initial description of the project was the following: "The number of attended outpatients is increasing since 2013, stressing out the scheduling of the Outpatient Pharmaceutical Care Unit agenda. Over 50% of patients who daily attended had not

been scheduled for that day. This results in waiting times of nearly an hour. Moreover, the stress of this systematic work overload may be affecting to the attention time, generating differences between attending staff and thus, an undesired variability on the caring process".

Measure Phase

During this phase, there were two data collecting processes going on.

- The first one was the *daily agenda* data (N = 125, K = 5). This data showed an outlook of the daily activity in the Outpatient Pharmaceutical Care Unit: number of scheduled visits, number of recorded visits at the end of the day, and number of missed visits. Each one of these metrics was shown globally (accounting for all patients) and split by visit type: first, successive and dispensing visit. The difference between successive and dispensing visits is that the former require the attention of pharmaceutical staff given that they may involve changes in medication doses or prescriptions, whereas dispensing could be performed both by pharmaceutical and by nursery staff.
- The second one was the outpatient visits data (N = 664, K = 13, L = 2). This database was designed on purpose by the Six Sigma technical team. It recorded, each day for two weeks, information about each one of the outpatients visiting the Outpatient Pharmaceutical Care Unit. This required the assistance of additional personal for the data collection and a strong engagement of all the staff, who responded very well to the demands of the technical team. The confection of a Fishbone diagram was used to determine potential causes affecting the waiting time. The included variables collected information about several aspects of the visit:

- Information about the patient: assigned clinical service (*service*), the hour of arrival to the desk (*arrival*) and the hour of start and end of the pharmaceutical care consultation (*enter, exit*).
- Information about the visit context: type of visit (*type*), day of visit (*date* and *weekday*) and turn of visit (*turn*). This last variable was obtained splitting the arrival hour into three categorical variables: turn 1 (from 8:00 a.m. to 10:30 a.m.), turn 2 (from 10:30 a.m. to 12:30 p.m.) and turn 3 (from 12:30 p.m. to 2:30 p.m.). This division corresponds to staff's experience about hour gaps in which the dynamics of the unit changes. Since these dynamics were probably non-linear (i.e., not necessarily that the later the visit, the longer the waiting time), we considered that the discretization of the hour into three categories might help us to spot bottlenecks and rush hours.
- Information about the treatment: if they were stored in the refrigerator (*refrigerator*), how many units were prescribed (*number*) and the route of administration (*via*).
- Information about the pharmacy unit staff who was attending the patient (*professional*) and the profile of the attending person (*profile*).

From all these variables, two output variables (i.e., critical-to-quality characteristics, CTQs or CQCs) were calculated: waiting time and attention time. The waiting time was computed as the difference in minutes between the enter and the arrival hours. The attention time was calculated as the difference in minutes between the exit and the enter hours.

The data were validated after checking the existence of transcription errors (such as negative durations). Since the pharmacy staff had their notes and records about the

visits, some of the errors could be solved, but all those entries with misleading information that could not be contrasted were not considered for further analysis.

The reference values were an average waiting time of 24,17 minutes and an average attention time of 4,78 minutes. These values were obtained from the data recorded in 2018. A descriptive summary of the initial situation from 2018 can be found in Supplementary Figure 4. This figure summarizes the CQCs (along with the overload of patients) at the beginning of the Six Sigma project. Setting reference values is key to quantitatively proving the usefulness and success of the improvement actions and, overall, of the Six Sigma project. Thus, reference values shown in Supplementary Figure 4, will be the base for the later comparison between the pre-Six Sigma project situation of the process (from 2018) and the post-Six Sigma situation (from data recorded in 2019).

Analyze Phase

In this stage, the goal was to establish factors affecting the CQCs: the waiting time and the attention time. Regarding the waiting time, the focus of the analysis was to check if there was any pattern of visits related to longer waiting times. The attention time presented another casuistic. Given the comments of patients arguing unfair differences in the caring process, the goal here was to establish if, for the same visits profile, there was an undesired variability in attention time. This would be reflecting a difference in the attention protocol followed by different members of the staff, which could impact the quality of the caring process. Supplementary Tables 1 and 2 contain parameters for the CQCs before and after the Six Sigma project.

To get this information, a PLS model was fitted, including all predictor variables and both CQCs. It is important to mention that, given the asymmetric distribution of the waiting and attention times, logarithmic versions of the aforementioned variables were

used. This analysis would let to identify the sources of variability of the attention process affecting each one of the CQCs. To interpret the relationships between process variables and CQCs found by the model, weighting plots were used.

The PLS analysis on the outpatient visits dataset of 2018 pointed out some interesting facts. In the weighting plot (Figure 3) the attention and waiting times (CQCs) are represented by red squares along with two almost orthogonal directions of variability, showing a lack of relationship between those two CQCs. The directions of variability aligned with the CQCs (red dashed lines) give information on the degree of correlation between predictors (i.e., process variables) and each one of the CQCs. The closer to the extreme of the red dashed line a predictor is, the more correlated it will be with the CQC associated with this direction of variability. This correlation will be positive if the predictor is located on the same side of the CQC, and negative if it is located on the opposite side of the red dashed line.

[Figure 3 here]

Figure 3. Weighting plot highlighting the relationships of process variables to the waiting time and the attention time.

This plot gives a clear picture of the latent structure of the process in the hospital pharmacy unit, showing that the process affecting the attention time is nearly independent of the process affecting the waiting time. The following analyses will focus on each CQC independently to ease the understanding of both processes.

Waiting time

The weighting plot in Figure 4 (a) shows that successive visits, Oncology or Haematology patient visits, Turn 3 visits and visits attended by a Resident are related to longer waiting times. On the contrary, dispensing visits and visits occurring in Turn 1 are associated with shorter waiting times.

[Figure 4 here]

Figure 4. (a) Weighting plot highlighting the relationships of process variables with the waiting time. Process variables positively correlated with waiting time are circled by the orange dotted contour and negatively correlated predictors are contained within the blue dashed contour. (b) PLS Coefficients plot for the relationship between variables in **X** and waiting time.

This information is also displayed in Figure 4 (b), where the PLS coefficients indicate the relationship between predictors (***X***) and the response variables (***Y***), in this case, waiting time. In this plot only statistically significant predictors (whose 95% jackknife confidence intervals do not contain the zero value) are shown. The interpretation of this plot is the following: process variables with positive and statistically significant ***B*** coefficients are positively correlated to waiting time, while process variables with negative and statistically significant ***B*** coefficients are negatively correlated to waiting time.

According to PLS results from Figure 4, it would be worth checking closely the relationship between the turn, the type of visit, the oncological and haematological specialities, and the resident profile concerning the waiting time. Figure 5 (a) shows an increasing trend of the waiting time along with the visit turns. This is particularly evident for successive visits. Besides, more than 80% of successive visits during turn 3 were for patients from the oncology (52,8 %) or haematology (31,9 %) service, as highlighted in Figure 5 (b).

[Figure 5 here]

Figure 5. (a) Boxplots of waiting time for each type of visit (dispensing, successive and first visit) and at each turn (T1, T2 and T3). (b) Pie chart of the types of assigned hospital service of successive visits during turn 3.

The association between these process variables relied on the fact that oncological and haematological visits were being attended mostly by a resident and only in Turn 3. Arriving at complex associations like this one can be tricky and time-consuming via univariate descriptive charts and analyses, whereas this relationship between several factors (onco/haema services, turn 3, successive visits and resident profile) and the waiting time, clearly stands out from the PLS analysis (Figure 4).

All this evidence was pointing towards a bottleneck associated with Turn 3 and Oncology/Haematology associated patients, as the PLS revealed. Moreover, in an eyeshot, the weighting plot from the PLS analysis also showed that visits scheduled at the first hour (Turn 1) or on Mondays, seemed to be related to shorter waiting times (Figure 4 a).

To quantify the statistical significance of these effects through the classical Six Sigma statistical toolkit, we run some ANOVA tests to double-check these hypotheses. A univariate ANOVA test finally confirmed the statistically significant effect of assigned hospital service on waiting time (p-value < 0.05). A Fisher LSD test with a 95% confidence level for multiple comparisons (Figure 6, b), shows that Oncological and Haematological visit profiles had statistically significant longer average waiting times than the rest of hospital services.

[Figure 6 here]

Figure 6. (a) 95% confidence intervals for the mean waiting time (minutes) for each assigned hospital service. (b) Least Significant Difference (LSD) analysis was performed on the waiting time for the different assigned hospital services.

Attention time.

The PLS weighting plot (Figure 7 a) and the PLS coefficients plot (Figure 7 b) showed that the nursery staff profile was associated with the shortest attention times, whereas pharmacists and resident profiles, were associated with the longest attention times.

[Figure 7 here]

Figure 7. (a) PLS weighting plot highlighting the relationship of process variables associated with the attention time. Process variables positively correlated with the attention time are circled by the orange dotted contour and negatively correlated predictors are contained within the blue dashed contour. (b) PLS coefficients for the relationship with variables in $\boldsymbol{X}$ and the attention time.

In a similar way to the waiting time analysis, a univariate doublecheck was carried out. An ANOVA test confirmed the statistically significant relationship between the attention time and the professional profile of the attending staff. To make a fair comparison, only those visits that could be attended by all professional profiles (and hence, were comparable) were included.

[Figure 8 here].

Figure 8. (a) 95% confidence intervals for the mean attention time for each professional profile. (b) Least Significant Difference (LSD) analysis was performed on the attention time for the different assigned staff profiles.

As it could be seen, nursery staff was showing statistically significant shorter attention times than the other professional profiles staff (Figure 8). These differences were, on average of two minutes. Considering that the average attention time of these visits was between 4-5 minutes, these differences represented between 40% and 50% of the visit duration. This variability could imply substantial differences in the attention procedure protocol. On one hand, if shorter times do not imply worse attention, then time was being wasted by longer attention procedures. But on the other hand, shorter times could be implying less careful attention, which could become critical in matters of health, such as this one is.

Improve Phase

After the analysis performed on waiting and attention times, the technical team of the Six Sigma project had a meeting to discuss the reported results and possible improvement actions.

Regarding the longer waiting times for the Onco-Haema visits, all the team agreed that attending all these visits on a specific Onco-Haema turn had become a bottleneck. This was initially done because these patients may change more frequently their medication, and that would need supervision and approval from a pharmacist specialized in oncological and haematological treatments. However, the distribution of the attention time for Onco-Haema visits (Figure 9) showed that most of them had a duration below 5 minutes. This meant that most of these visits were not needing a comprehensive re-evaluation of the medication and were just for drug dispensing.

[Figure 9 here]

Figure 9. Histogram of the attention time for oncological and haematological visits.

To alleviate this bottleneck, it was proposed that those patients whose oncologist had not changed the medication, did not need a specific visit. Thus, they could be attended by pharmacists all morning, and not only during turn 3.

Another improvement to reduce the waiting time was to add a specific hour for the outpatient scheduling. The usual procedure involved only a day-of schedule. However, Turn 1 on the morning had shown minor waiting times (see Figure 5 a), which indicated that scheduling more patients at this time of the day would improve the patient flow, preventing the accumulation on Turn 3.

Finally, there were some improvements as well regarding the variability in the attention time. Attention protocols were designed and implemented to standardize the time and the depth of the attention for each visit.

All the proposed changes were implemented in November 2018. In May 2019, the Outpatient Pharmaceutical Care Unit had implemented regularly all the proposed changes.

Control Phase

Once the improvements were shown to work, the tracking of the activity was kept on. An intensive data collection for another two weeks was done to evaluate the effects of the changes in the unit workflow. This data collection yielded two new datasets: the daily agenda data of 2019 (N = 124, K = 5) and the outpatients visits data of 2019 (N = 1043, K = 13, L = 2). The comparison between the initial and final output values can be seen in Figure 10 and Supplementary Tables 1 and 2.

[Figure 10 here]

Figure 10. 95 % confidence intervals comparing the situation before (2018) and after (2019) the changes in the Outpatients Pharmaceutical Care Unit for the overload on

mornings (a), the waiting time for oncological and haematological patients (b) and for the attention time for the three staff profiles (c).

As it can be seen in Figure 10 (a) and in Supplementary Figure 5, the overload of patients changes from its historical values (Jan – October 2018), gradually decaying over November 2018 to March 2019, and finally stabilizing around April 2019. These differences are stated every month (Supplementary Figure 6). Table 1 shows the LSD interval that confirmed these differences to be statistically significant (p-value<0.05): in 2019 there were on average nearly 20 patients less of daily overload than in 2018.

[Table 1 here]

Table 1. Fisher LSD interval for the difference between mean outpatients' overload of 2019 and 2018.

Regarding the waiting time, there was a significant reduction in the overall waiting time, between 3 and 7 minutes per patient. This meant a reduction of the mean waiting time from 24 minutes per patient to 19 minutes (Supplementary Figure 7), achieving the project goal (with one minute more of reduction). This difference was even more noticeable for the waiting time for the Oncological and Haematological patients (Figure 10 b). Table 2 shows the 95% LSD confidence intervals for the difference between average waiting times for these two services, where a statistically significant reduction can be appreciated. Moreover, there was not a statistically significant increase in the waiting time for all other medical specialities (95% LSD interval contains the zero value). This result provided a solid improvement and achievement of the 6S project.

[Table 2 here]

Table 2. Fisher LSD intervals for the differences between mean waiting times of 2019 and 2018, for oncological and haematological outpatients and all other medical specialities.

Finally, Figure 10 (c) shows that differences between attention times were also reduced. Table 3 shows two interesting things. First, that the biggest time gap between attention times is now 1.2 minutes/patient, which was a reduction of 52% concerning the previous maximal difference (2.5 min/patient, Supplementary Table 3). Secondly, differences are presented for a different professional profile (resident) after the protocols update. This outcome can be used as evidence for future updates of the attention protocol, focusing now on reducing the variability on the attention time due to the different performance of professional profiles.

[Table 3 here]

Table 3. Fisher LSD intervals for the differences between mean attention times of 2019, for different professional profiles.

Table 4 summarizes the goals, the implemented changes and the outcomes obtained for each one of the improvement goals. A more detailed description of the changes can be found in Supplementary Table 4.

[Table 4 here]

Table 4. Summary of the improvement goals, the implemented changes on the workflow of the hospital pharmacy unit and the results obtained after the implementation.

After controlling the improvements, the collection of the agenda database was kept on, registering all the daily information about outpatient schedules. Now, this database provides a continuous flow of information that is analyzed by the staff of the Outpatient Pharmaceutical Care Unit and the Department of Pharmacy.

The success of this project also meant the configuration of a solid continuous improvement technical team in the Outpatient Pharmaceutical Care Unit, who is responsible for future updates and changes in response to the results of this project and to further data.

Conclusions

In this work, PLS has been incorporated into the Six Sigma toolbox to explore and analyze the dataset from a Six Sigma project in the Outpatient Pharmaceutical Care Unit in the Department of Pharmacy at a university hospital. In contrast to univariate techniques, PLS shows in a single shot an overall picture of how input and output variables of the caring process are correlated, providing a clear interpretation of the results that becomes crucial for process understanding and implementing actions for improvement. This is a clear added value of using PLS in Six Sigma projects, no matter the number of X´s and Y´s. Even when the number of variables is not so high, as in this case study, the proposed two-step procedure clearly showed how PLS findings efficiently guide the confirming process by using classical Six Sigma tools such as ANOVA, simplifying the number of statistical tests to be carried out, if needed.

Thus, the classical Six Sigma DMAIC scheme can be upgraded for a more effective and time-saving methodology able to work with increasingly complex databases, by including latent variable-based techniques, such as PLS, going forward to the next generation of process improvement methodology in 4.0 environments: the Multivariate Six Sigma.

Acknowledgements

The authors acknowledge the work of the staff of the Outpatient Pharmaceutical Care Unit at the Department of Pharmacy of Hospital Universitario Doctor Peset (Valencia,

Spain). This work was supported by the Universitat Politècnica de València under the program PAID-01-17 and the Valencian regional government research grant:AICO/2021/111, INDOPT4.0.

References

[1] S. L. Furterer, "Applying Lean Six Sigma methods to reduce length of stay in a hospital's emergency department," *Qual. Eng.*, vol. 30, no. 3, pp. 389–404, Jul. 2018, doi: 10.1080/08982112.2018.1464657.

[2] A. C. Honda, V. Z. Bernardo, M. C. Gerolamo, and M. M. Davis, "How lean six sigma principles improve hospital performance," *Qual. Manag. J.*, vol. 25, no. 2, pp. 70–82, 2018, doi: 10.1080/10686967.2018.1436349.

[3] D. Tlapa *et al.*, "Effects of Lean Healthcare on Patient Flow: A Systematic Review," *Value Heal.*, vol. 23, no. 2, pp. 260–273, 2020, doi: 10.1016/j.jval.2019.11.002.

[4] I. Font Noguera, M. J. Fernández Megía, A. J. Ferrer Riquelme, S. Balasch I Parisi, M. D. Edo Solsona, and J. L. Poveda Andres, "Mejora del proceso farmacoterapéutico del paciente hospitalizado mediante la metodología Lean Seis Sigma," *Rev. Calid. Asist.*, vol. 28, no. 6, pp. 370–380, 2013, doi: 10.1016/j.cali.2013.04.003.

[5] J. P. Womack and D. T. Jones, "Lean thinking—banish waste and create wealth in your corporation," *J. Oper. Res. Soc.*, vol. 48, no. 11, p. 1148, 1997.

[6] K. Linderman, R. G. Schroeder, S. Zaheer, and A. S. Choo, "Six Sigma: a goal-theoretic perspective," 2003. doi: 10.1016/S0272-6963(02)00087-6.

[7] M. J. Harry, P. S. Mann, O. C. De Hodgins, R. L. Hulbert, and C. J. Lacke, *Practitioner's guide to statistics and lean six sigma for process improvements*. John Wiley & Sons, 2010.

[8] G. Improta *et al.*, "Lean thinking to improve emergency department throughput at AORN Cardarelli hospital," *BMC Health Serv. Res.*, vol. 18, no. 1, pp. 1–9, 2018, doi: 10.1186/s12913-018-3654-0.

[9] J. Robinson, M. Porter, Y. Montalvo, and C. J. Peden, "Losing the wait: improving patient cycle time in primary care," *BMJ open Qual.*, vol. 9, no. 2, 2020, doi: 10.1136/bmjoq-2019-000910.

[10] W. M. Ma, H. Zhang, and N. L. Wang, "Improving outpatient satisfaction by extending expected waiting time," *BMC Health Serv. Res.*, vol. 19, no. 1, pp. 1–7, 2019, doi: 10.1186/s12913-019-4408-3.

[11] A. Ferrer-Riquelme, "Multivariate Six Sigma: a key improvement strategy in Industry 4.0," 2021 *(accepted Qual. Eng)*.

[12] P. Geladi and B. R. Kowalski, "Partial Least-Squares Regression: a Tutorial," *Anal. Chim. Acta Elsevier Sci. Publ. B.V*, vol. 185, pp. 1–17, 1986, doi: 10.1016/0003-2670(86)80028-9.

[13] A. Höskuldsson, "PLS regression.," *J. Chemom.*, vol. 2, no. August 1987, pp. 581–591, 1988, doi: 10.1002/cem.1180020306.

[14] C. M. Jaeckle and J. F. MacGregor, "Industrial applications of product design through the inversion of latent variable models," *Chemom. Intell. Lab. Syst.*, vol. 50, no. 2, pp. 199–210, 2000, doi: https://doi.org/10.1016/S0169-7439(99)00058-1.

[15] A. Ismail *et al.*, "DMAIC Six Sigma Methodology in Petroleum Hydrocarbon Oil Classification," *Int. J. Eng. Technol. Vol 7, No 3.14 Spec. Issue 14*, 2018, doi: 10.14419/ijet.v7i3.14.16868.

[16] R. S. Peruchi, P. R. Junior, T. G. Brito, A. P. Paiva, P. P. Balestrassi, and L. M. M. Araújo, "Integrating Multivariate Statistical Analysis Into Six Sigma DMAIC

Projects: A Case Study on AISI 52100 Hardened Steel Turning," *IEEE Access*, vol. 8, pp. 34246–34255, 2020, doi: 10.1109/ACCESS.2020.2973172.

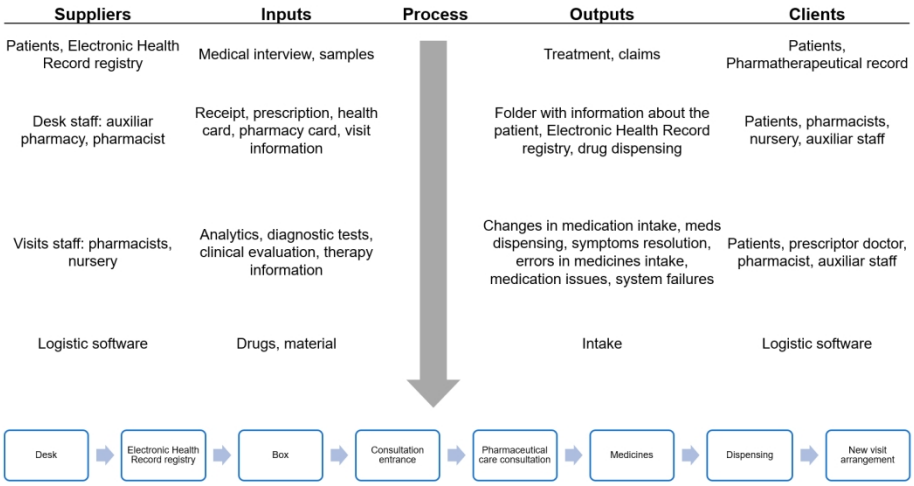
[17] D. Palací-López, J. Borràs-Ferrís, L. T. da Silva de Oliveria, and A. Ferrer, "Multivariate six sigma: A case study in industry 4.0," *Processes*, vol. 8, no. 9, pp. 1–20, 2020, doi: 10.3390/PR8091119.

[18] P. Grima, L. Marco-Almagro, S. Santiago, and X. Tort-Martorell, "Six Sigma: hints from practice to overcome difficulties," *Total Qual. Manag. Bus. Excell.*, vol. 25, no. 3–4, pp. 198–208, Feb. 2014, doi: 10.1080/14783363.2013.825101.

[19] M. Barker and W. Rayens, "Partial least squares for discrimination," *J. Chemom.*, vol. 17, no. 3, pp. 166–173, Mar. 2003, doi: 10.1002/cem.785.

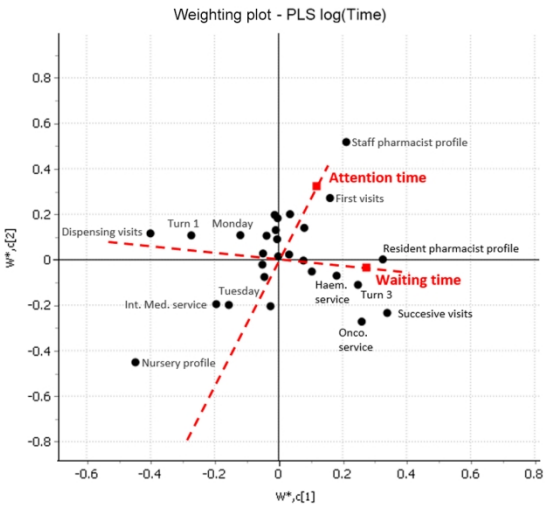
Timeline of the Six Sigma project, indicating the data recording periods and implementation of changes.

406x101mm (118 x 118 DPI)



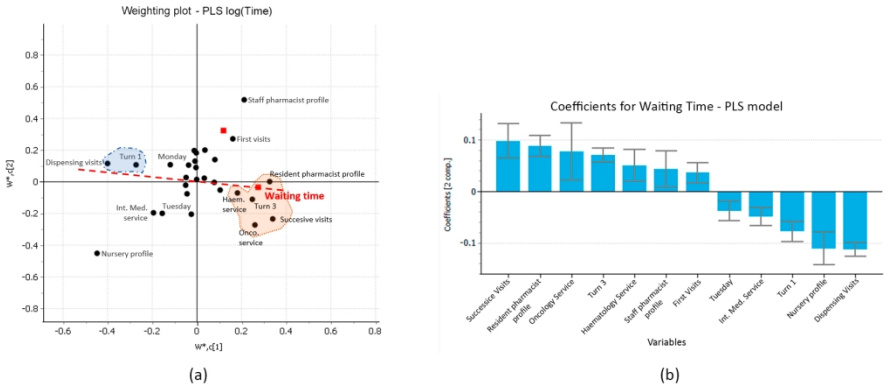
SIPOC diagram of the workflow in the Outpatient Pharmaceutical Care Unit.

406x226mm (118 x 118 DPI)



Weighting plot highlighting the relationships of process variables to the waiting time and the attention time.

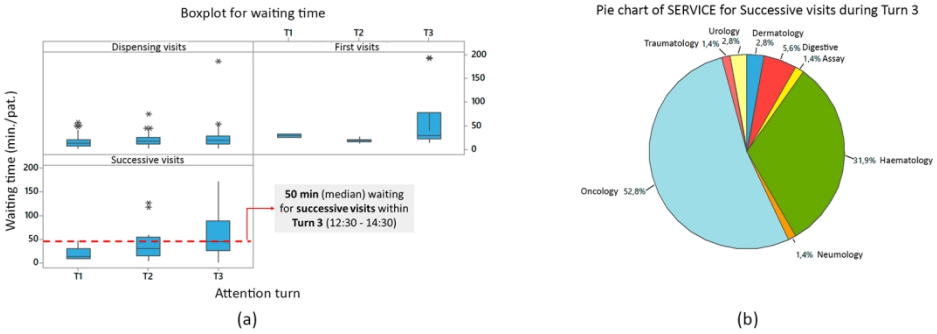
406x220mm (118 x 118 DPI)



(a) (b)

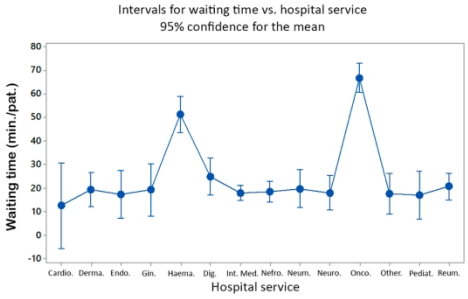
(a) Weighting plot highlighting the relationships of process variables with the waiting time. Process variables positively correlated with waiting time are circled by the orange dotted contour and negatively correlated predictors are contained within the blue dashed contour. (b) PLS Coefficients plot for the relationship between variables in **X** and waiting time.

406x171mm (118 x 118 DPI)



(a) Boxplots of waiting time for each type of visit (dispensing, successive and first visit) and at each turn (T1, T2 and T3). (b) Pie chart of the types of assigned hospital service of successive visits during turn 3.

374x135mm (118 x 118 DPI)



(a)

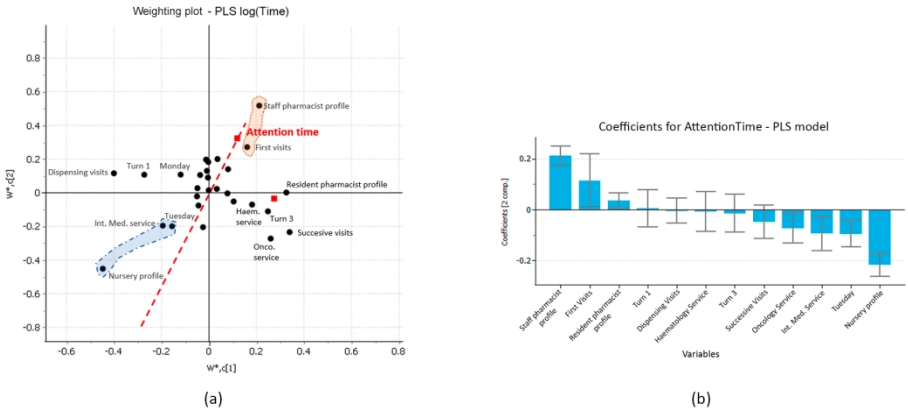
Fisher LSD with confidence level of 95%

Service	N	Average	Group		
Onco	51	66,88	A		
Haem	34	51,29		B	
Dig.	32	25,13			C
Reu.	61	20,90			C
Neum.	31	19,87			C
Derma	37	19,38			C
Gin.	16	19,38			C
Nefro.	98	18,61			C
Neuro.	36	18,17			C
Int.Med.	198	18,11			C
Other	26	17,73			C
Endo.	19	17,58			C
Pediat.	19	17,11			C
Cardio	6	12,67			C

(b)

(a) 95% confidence intervals for the mean waiting time (minutes) for each assigned hospital service. (b) Least Significant Difference (LSD) analysis was performed on the waiting time for the different assigned hospital services.

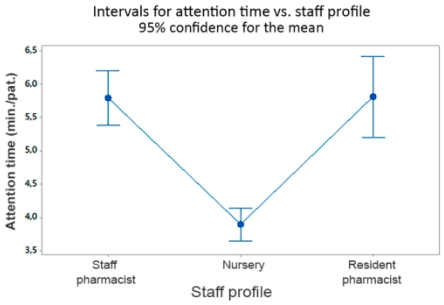
406x206mm (118 x 118 DPI)



(a) (b)

(a) PLS weighting plot highlighting the relationship of process variables associated with the attention time. Process variables positively correlated with the attention time are circled by the orange dotted contour and negatively correlated predictors are contained within the blue dashed contour. (b) PLS coefficients for the relationship with variables in **X** and the attention time.

406x184mm (118 x 118 DPI)



(a)

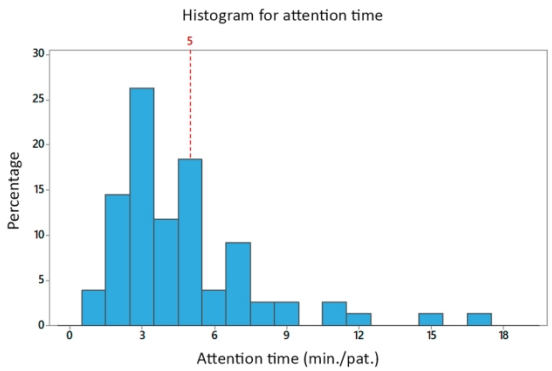
Fisher LSD with confidence level of 95%

Staff	N	Average	Group	
Staff pharmacist	128	5,797	A	
Nursery	344	3,898		B
Resident pharmacist	58	5,810	A	

(b)

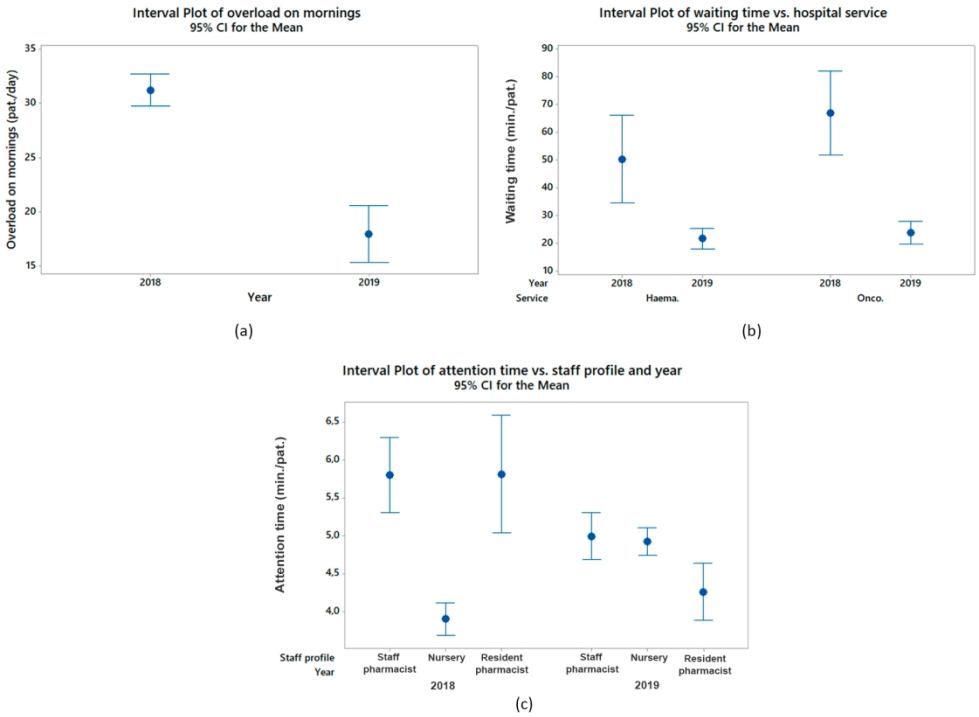
(a) 95% confidence intervals for the mean attention time for each professional profile. (b) Least Significant Difference (LSD) analysis was performed on the attention time for the different assigned staff profiles.

334x120mm (118 x 118 DPI)



Histogram of the attention time for oncological and haematological visits.

406x156mm (118 x 118 DPI)



95 % confidence intervals comparing the situation before (2018) and after (2019) the changes in the Outpatients Pharmaceutical Care Unit for the overload on mornings (a), the waiting time for oncological and haematological patients (b) and for the attention time for the three staff profiles (c).

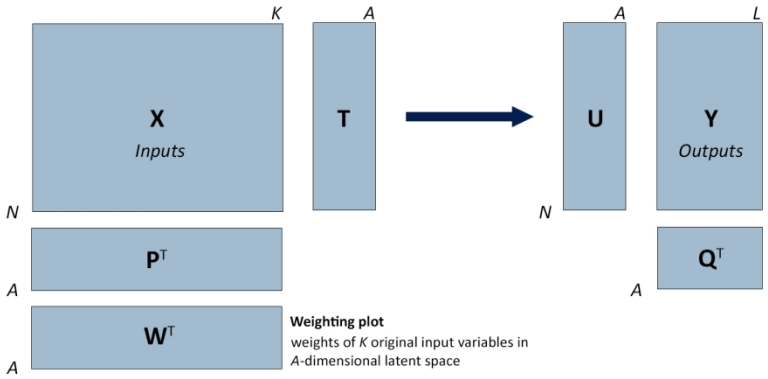
406x305mm (118 x 118 DPI)

Metric	LSD intervals at 95% for the difference 2019 – 2018 (patients/day)
Overload.	[-21,38; -15,23]

Service	LSD intervals at 95% for the difference 2019 – 2018 (min./patient)
Onco.	[-56,45; -29,85]
Haema.	[-39,55; -17,78]
Others.	[-1,71; 1,09]

Profiles	LSD intervals at 95% for differences between staff profiles (min./patient)
Nurs – Staff phar.	[-0,42; 0,27]
Res. Phar. – Staff phar.	[-1,21; -0,26]
Res. Phar. – Nurs.	[-1,09; -0,24]

Improvement goal	Implemented change	Outcome
Overload decrease	Schedule all patients with day and hour	Reduction of average overload between 15 and 21 patients/day:
To reduce waiting time of turn 3	Non-specific oncological and haematological visits are moved to turns 1 and 2	Reduction of waiting time between 3 and 7 min./patient.
To reduce differences between attention times for Scheduled Dispensing visits.	Standard caring protocols	Reduction from 2.5 min. of difference in 2018, to 1.2 min. in 2019.



Schematic representation of the PLS data matrices and coefficient matrices.

406x190mm (118 x 118 DPI)

Project Charter: Definition sheet

Date: July 2nd 2018

Title of the project	
Reduction of the average waiting time for patients of the hospitalary pharmacy unit of external patients	
Project leader BB:	**Process owner**
Ana Moya	Mónica Climente

Team members

Marta Hermenegildo, Irene Micciché, Ángel Marcos, Alba González, Tamara Lidia Paredes, Mercedes Riera, Carlos Cortés

Agents involved

Champion: Mónica Climente
Black Belt: Ana Moya
Work team Tamara L.P., Irene M., Mercedes R., Ángel M., Carlos C., Marta H., Alba G.

Problem description

The number of attended patients is growing since 2013, which generates planning difficulties in the agenda of the unit, dealing everyday with a 50% of non-scheduled patients. This generates waiting times of almost one hour. Moreover, the stress generated by this overload seems to be affecting also to the attention time, causing differences on the attention procedure between staff members.

Goals:	Metrics:	Initial value	Target value
Reduction of the average waiting time between 8:30 and 14:30.	Minutes/patient	24	20
Reduction of the overload of visits between 8:30 and 14:30.	Patients/day	28	14

Expected economic results

(Results not measured in ecomomic terms)
Process improvement in terms of eliminating non-quality from the attention service provided to patients.
Process improvement increasing the efficiency of the service and reducing bottlenecks of the process.

Expected benefits for the clients

Increased satisfaction from the external client perspective, thanks to the reduction of the waiting time.
Increased satisfaction of the internal client (hospital pharmacy unit staff), thanks to the reduction of the overload and the pressure in the work environment.

Available resources

Staff: hospital pharmacy unit personnel (two residents, two pharmaceuticals, one nurse and two auxiliary staff.

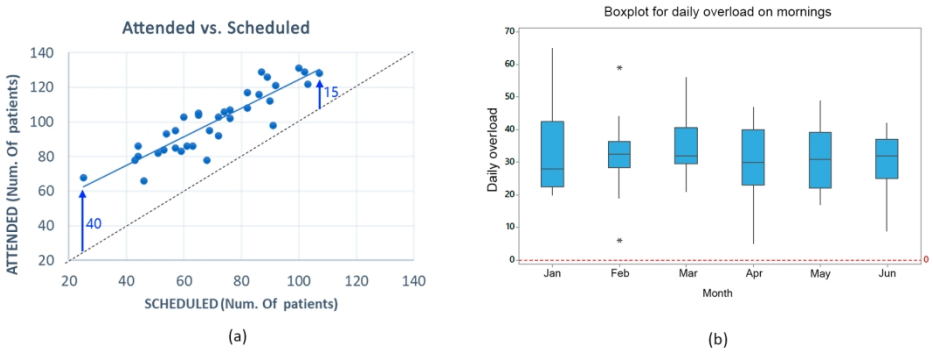
Material: laptop and software for the statistical analysis of the data.

Project constraints

Starting date	July 2nd of 2018	Expected ending date:	October 5th of 2018

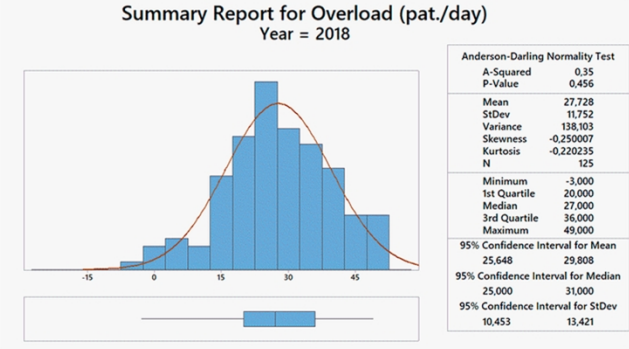
Project Charter of the Project.

406x622mm (118 x 118 DPI)

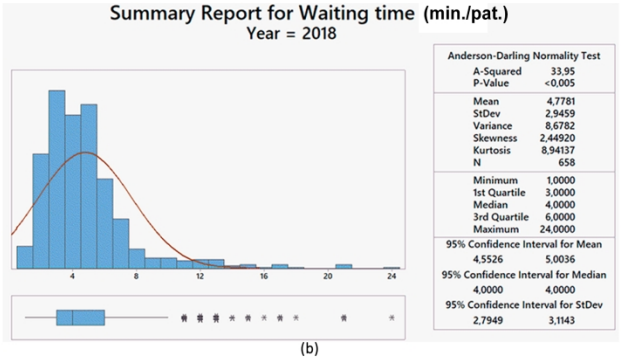


(a) Attended versus scheduled patients during 2018, showing a systematic overload appreciated by the consistent position of data points above the diagonal representing the equivalence between the planned and the recorded visits for each day. (b) Boxplot of daily overload during the data recorded in 2018 (January to June), where the overload boxplots are above the zero reference (red dashed line) line for every month analyzed.

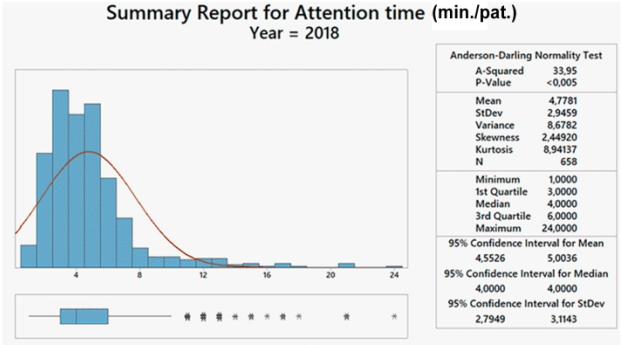
406x158mm (118 x 118 DPI)



(a)



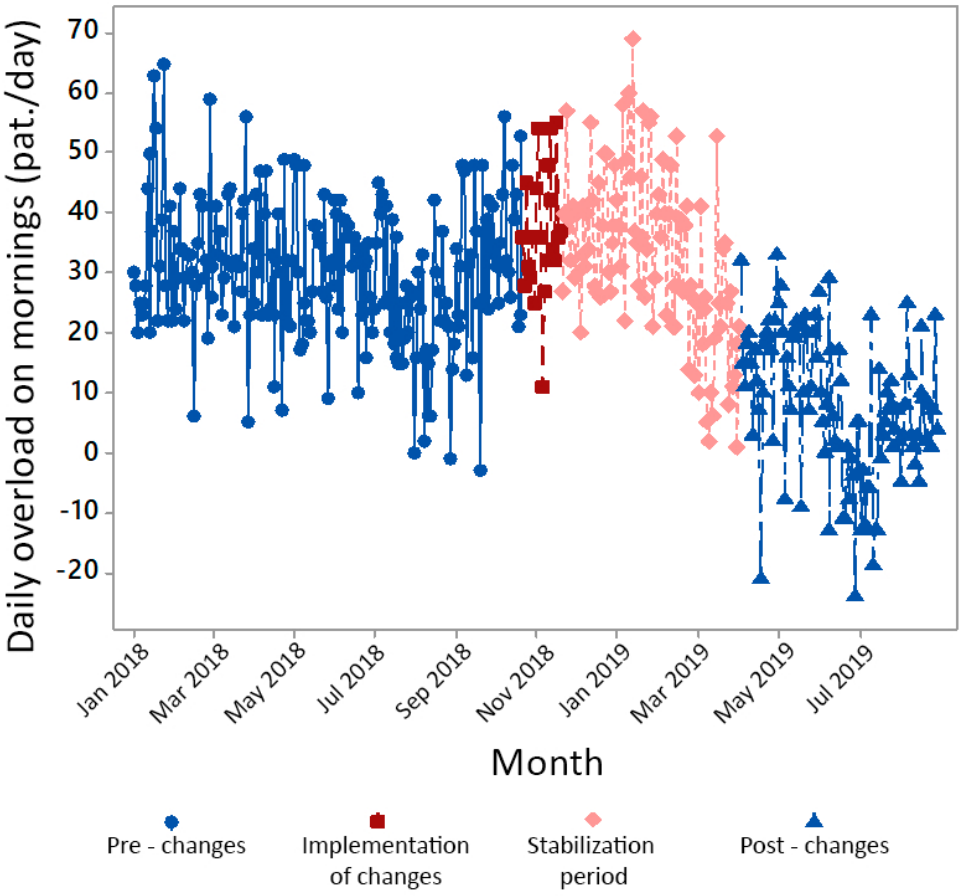
(b)



(c)

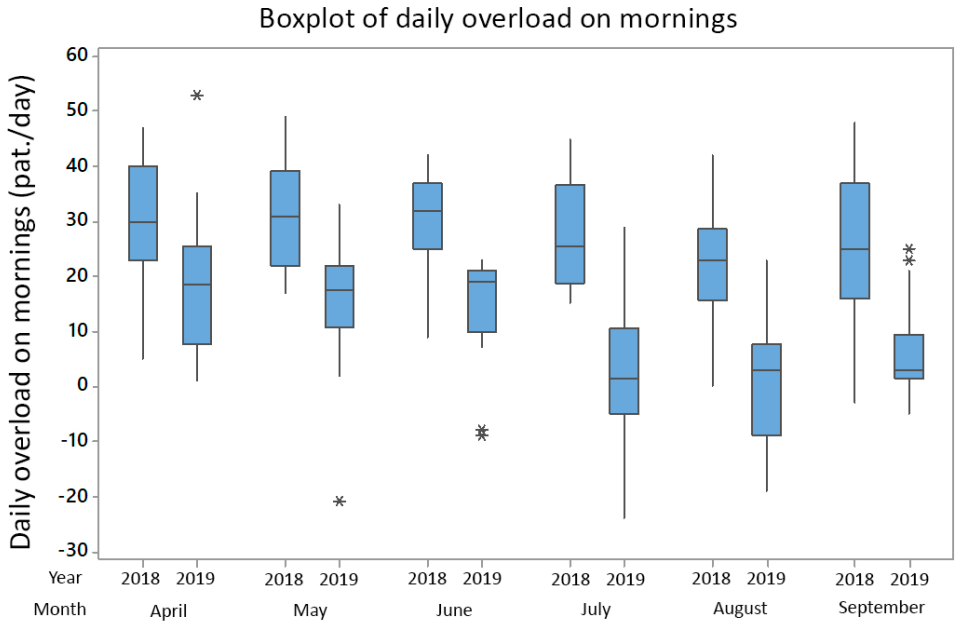
Descriptive statistics for the daily overload of patients (a), the waiting time (b) and the attention time (c) according to the data from 2018.

406x724mm (118 x 118 DPI)



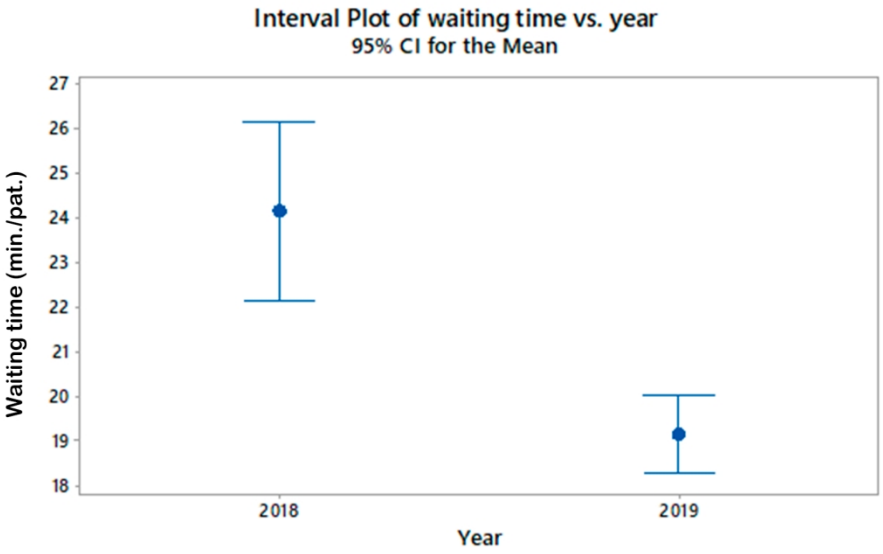
Time series of the overload from January 2018 to September 2019.

173x173mm (118 x 118 DPI)



Boxplots of the overload over the comparable months (April to September) of 2018 and 2019.

231x153mm (118 x 118 DPI)



Confidence Intervals for the waiting of 2018 and 2019.

406x305mm (118 x 118 DPI)

Output	Average (I.C.$_{95\%}$) 2018	Average (I.C.$_{95\%}$) 2019
Overload	27,73 (25,56; 29,90) pat. /day	9,42 (7,24; 11,61) pat. /day
Waiting time	24,17 (22,17; 26,17) min. /pat.	19,17 (18,30; 20,04) min. /pat.
Attention time	4,78 (4,55; 5,00) min. /pat.	5,45 (5,28; 5,62) min. /pat.

Output	Median 2018	Median 2019	[Q1; Q3] 2018	[Q1; Q3] 2019
Waiting time	18 min. /pat.	17 min. /pat.	[10; 28] min. /pat.	[7; 28] min. /pat.
Attention time	4 min. /pat.	5 min. /pat.	[3; 6] min. /pat.	[4; 6] min. /pat.

Staff	LSD intervals at 95% for the differences between staff profiles (min./patient)	
Nurs – Staff phar.	-2,38	-1,42
Res. Phar. – Staff phar.	-0,72	0,74
Res. Phar. – Nurs.	1,26	2,57

Improvement goal	Implemented change	Outcome
Overload decrease	Schedule all patients with day and hour	Average overload decrease (15; 21) patients/day: • Consistent reduction in all months. • Average reduction in Dispensing visits (13 ; 17) patients/day. • and Successive visits (6 ; 10) patients/day. • Average increase for First Visits(4 ; 5) patients/day.
To reduce waiting time of turn 3	Non-specific oncological and haematological visits are moved to turns 1 and 3	Reduction of waiting time between 3 y 7 min./patient. • By hospital services: reduction in Onco. (30 ; 56) min./patient, and in Haema. (18 ; 40) min./patient - without longer waiting times for the rest of medical specialties. • By turns: average decrease in turn 3 (10 ; 18) min./patient and increase in turn 2 (0,15 ; 5) min./patient.
To reduce differences between attention times	Standard attention protocols	Reduction of differences between average attention time for Dispensing visits between staff: • Maximum differences for average attention time in 2018 were of 2,5 minutes; in 2019 of 1,2 minutes. • In 2019 visits attended by residents were faster on average than those attended by deputies (-1,21; -0,26 min./patient) and by nursery (-1,09; -0,24 min./patient).