ANALYSIS OF THE DRUG PRODUCT QUALITY PARAMETERS THROUGH STATISTICAL METHODS

Dipl. Mag. Toni Paneva, PhD student[1], Assos. Prof. Dr. Eng. Elena Koleva[1,2]

[1] University of Chemical Technology and Metallurgy, 8, Kliment Ohridski, Blvd. Sofia, 1756, Bulgaria
[2]Institute of Electronics – BAS, 72 Blvg. "Tsarigradsko shosee", 1784 Sofia, Bulgaria

[1]tony_pan@abv.bg
[2]eligeorg@abv.bg

Abstract: *Correlations between some basic quality parameters of a drug product, which are very important for the pharmaceutical manufacture of solid dosage forms are studied, on the base of experimental data from 135 batches. Factor analysis, control charts and histograms are implemented for the quality parameters - strength of tablet, assay of the active substance, acid resistance, disintegration), water content, drug dissolution after 150 minutes and drug dissolution after 165 minutes.. The parameter drug dissolution after 120 minutes is dropped from the analysis, because it was observed to have a zero value for all measurements. The main aim of this article is to search for a possible underlying structure in the variables with exploratory factor analysis, to monitor whether the quality parameters' variation is consistent and the empirical distributions of the most important quality parameters using control charts and histograms.*

Keywords: QUALITY, PRINCIPAL COMPONENTS ANALYSIS, CONTROL CHARTS, HISTOGRAM, PHARMACEUTICAL MANUFACTURE OF SOLID DOSAGE FORMS

1. Introduction

That part of the Quality Management, which ensures that the products are consistently produced and controlled to the quality standards appropriate to their intended use and as required by the Marketing Authorization, Clinical Trial Authorization or product specification, is the Good Manufacturing Practice. It is connected with both production and quality control. Quality control includes sampling, specifications and testing, and with the organization, documentation and release procedures which ensure that the necessary and relevant tests are actually carried out and that materials are not released for use, nor products released for sale or supply, until their quality has been judged to be satisfactory [1].

Various experimental and analytical techniques can be used to characterize the process. The purpose of such development is to understand the process as well as the connection of the input parameters with the output characteristics. The main features of a product and process are: Critical Quality Output Characteristics, Input Process Parameters, Process Capabilities, and Quality Control Techniques [2].

Design of Experiments, Response Surface Methodology (RSM) [3], Robust Engineering Design Methodology [4], combined with methods for parameter optimization; provide opportunities for improving the quality of the drug production, for obtaining cheaper production products through an appropriate choice of components, to save raw materials and energy, to ensure the sustainability of the quality indicators in terms of noise and variation impacts of different nature.

Historically, the drug manufacture has taken place in pharmacies, and only after the start of the economic revolution and the commencement of industrial production of goods, industrial drug production has been identified [5]. One of the main stages of the life cycle related to the provision of drugs is their production. First, the drug should be authorized for use and in order to reach the patient it must be manufactured and marketed. The manufacture of a drug product includes all operations related to the procurement of materials, their processing in the production, packaging and labeling, quality control, batch release, storage, dispatching and related controls.

The aim of the article to study the quality parameters – strength of tablet, quantitative content of the active substance, acid resistance, disintegration, water content, dissolution after 150 minutes and dissolution after 165 minutes on the base of experimental data from 135 batches. Factor Analysis a method for modeling observed variables, and their covariance structure, in terms of dimension reduction of the observed variables by a smaller number of underlying unobservable (latent) factors. It is implemented to investigate the correlations and to define the more significant for the variance of the measured quality characteristics. Statistical process control methods (control charts and histograms) are applied for the quality analysis of the process control stability and the empirical distribution of the quality characteristics. The experimental data are obtained from laboratory tests of a drug product.

2. Factor Analysis

Experimental data from 135 batches are obtained. The studied quality parameters are strength of tablet (Y_1), assay of the active substance (Y_2), acid resistance (Y_3), disintegration (Y_4), water content (Y_5), dissolution after 150 minutes (Y_6) and dissolution after 165 minutes (Y_7). The parameter dissolution after 120 minutes is dropped from the analysis, because it was observed to have a zero value for all measurements.

Factor analysis assumes that the covariation in the observed variables is due to the presence of one or more latent variables (factors) that exert causal influence on these observed variables. Principal components analysis method can be used for estimating the parameters of the factor model. The observed variables are modeled as linear functions of the latent factors. The goal is to explain the maximum amount of variance with the fewest number of latent factors. An oblique rotation, which is appropriate for both uncorrelated and correlated factors, is implemented on the factor solution for the purpose of making the solution easier to interpret [6]. Through the factor analysis, the data are interpreted in a new way.

2.1. Results and interpretation of the factor analysis

The eigenvalue results determine the number of applied latent factors. The eigenvalue show the contribution of the respective factor in explaining the total variation in the observed variables. According to the Kaiser criterion, retain principal components with eigenvalues greater than 1. The eigenvalues of the latent factors are presented on the scree plot on Fig. 1.

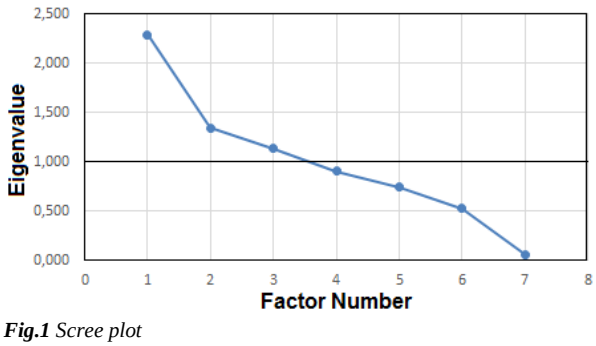

Fig.1 *Scree plot*

The method "scree" determines the number of factors. The scree plot displays the factor number versus the corresponding eigenvalue. The factor analysis is performed on the correlation matrix (Table 1) because the variables are standardized (each variable has a variance of 1, and the total variance is equal to the number of variables used in the analysis – 7). Three factors are with eigenvalues (variances) greater than one, and it is suggested that they adequately explain the variation in the data. The percent of total variance accounted for each of these factors are correspondingly 32.668%, 19.175% and 16.251% (or totally 68.094%).

Table 1. Correlation matrix

	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
Y_1	1.000	0.018	-0.187	0.290	-0.171	0.058	0.032
Y_2	0.018	1.000	-0.387	0.041	-0.028	0.231	0.224
Y_3	-0.187	-0.387	1.000	-0.010	0.009	-0.268	-0.270
Y_4	0.290	0.041	-0.010	1.000	-0.041	0.187	0.105
Y_5	-0.171	-0.028	0.009	-0.041	1.000	0.068	0.069
Y_6	0.058	0.231	-0.268	0.187	0.068	1.000	0.937
Y_7	0.032	0.224	-0.270	0.105	0.069	0.937	1.000

In Table 2 the obtained factor matrix is presented for the extracted three significant factors. It contains the non-rotated factor loadings, which are the correlations between the variables and the factors. The problem with this analysis is that some of the variables are highlighted in more than one column.

Table 2. Factor matrix

	Factor		
	1	2	3
Y_6	**0.905**	-0.194	0.284
Y_7	**0.892**	-0.245	0.248
Y_1	0.205	**0.779**	0.089
Y_4	0.270	**0.554**	**0.476**
Y_5	0.045	**-0.551**	0.178
Y_2	**0.495**	0.042	**-0.628**
Y_3	**-0.556**	-0.160	**0.579**

The factor pattern matrix in Table 3 contains the coefficients for the linear combination of the variables. The applied rotation method is oblique rotation with Kaiser Normalization. The total amount of variation explained by the rotated factor model is the same, but the contributions are not the same from the individual factors. We gain a cleaner interpretation, but the first factor is not going to explain as much of the variation.

Table 3. Pattern matrix

	Factor		
	1	2	3
Y_6	**0.943**	0.060	-0.140
Y_7	**0.930**	-0.001	-0.161
Y_1	-0.025	**0.797**	-0.086
Y_4	0.284	**0.713**	0.253
Y_5	0.292	**-0.458**	0.190
Y_2	0.091	-0.065	**-0.792**
Y_3	-0.129	-0.072	**0.787**

The coefficients indicate the relative weight of each variable in the factor. The bigger the absolute value of the coefficient, the more important the corresponding variable is in constructing the factor. The interpretation of the factors is subjective and requires knowledge of the data:

- For Factor 1 is primarily a measure of the variables Y_6 and Y_7 - Dissolution after 150 minutes and Dissolution after 165 minutes. As one of the variables increases the other also increases.
- For Factor 2 contribution have the variables Y_1, Y_4 and Y_5: Strength of tablet, Disintegration and Water content. The Strength of tablet and the Disintegration increase with the decrease of the water content.
- For the Component 3 contribution have the variables Y_2 and Y_3: assay of the active substance (Y_2) and acid resistance (Y_3). With the increase of the Assay of active substance, the Acid resistance decreases.

3. Control charts and histograms

A control chart provides a simple way of visually tracking a process in order to identify the appearance of trends in time. It consists of a horizontal plot of an ongoing performance characteristic (or group measures like mean, range, standard deviation, etc.). A control chart always has a central line for the average, an upper line for the upper control limit and a lower line for the lower control limit. These lines are determined from historical data. By comparing current data to these lines, you can draw conclusions about whether the process variation is consistent (in control) or is unpredictable (out of control, affected by special causes of variation). Overlaid lines show evaluation criteria such as allowed tolerance limits. The control chart highlights poor quality by showing when a measurement lies outside the expected variation. It shows when a process is trending toward failure. The purpose of a control chart is to monitor common cause variation and to detect the special cause variation. The expectation for a process is that it is under statistical control. This means that the only component of variation is the measured noise.

A typical control chart has control limits set at values such that if the process is under control, all points will be between the upper control limit (UCL) and the lower control limit (LCL).

Individual control charts (or Shewhart control chart) are used whenever the sample size for process monitoring is one observation per batch [6].

The quality of the drug product is determined by how much the quality parameter matches the desired value and how large the deviations from it are. The histogram studies the empirical distribution of the observations. The histogram is a bar chart that shows how often one or another value of the analyzed parameter appears.

For the three most significant quality parameters, that took part as variables with the highest weights in the factor analysis, control charts and histograms are made.

3.1. Control chart and histogram for the quality parameter Strength of the tablet (Y_1).

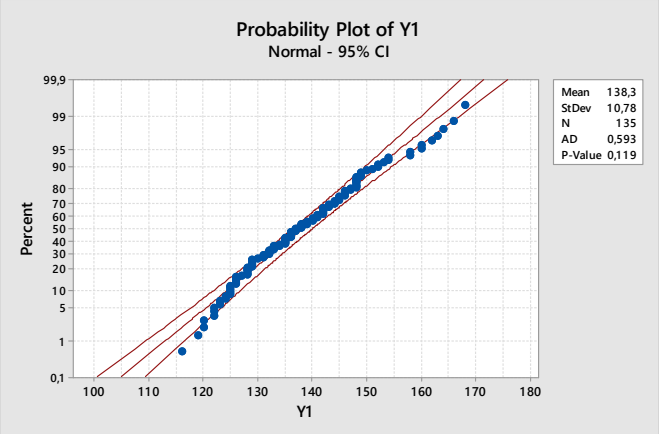

Fig. 2. *Normality test*

In order to analyze the data with control charts and histograms, it is necessary to perform a normality test. Anderson-Darling test is implemented and the results are presented in Fig. 2. According them, the quality parameter Strength of tablet has normal distribution.

The individual values control chart for Strength of tablet is presented in Fig. 3. The control limits are wide. There are three point outside the control limits (points 30, 31 and 66). The data are out of statistical control. That means, that after several manufactured batches it is possible to have measurements that are not in specification limits. It is necessary to take measures to eliminate the reasons for this process behavior, since there are special causes for such variation. The histogram shows the data distribution (Fig.4). It can be seen that the distribution seems to be bimodal and there is asymmetry. The natural tolerance limits can be determined and compared with the specification limits and further the Process Capability Index and the Process Performance Index can be calculated.

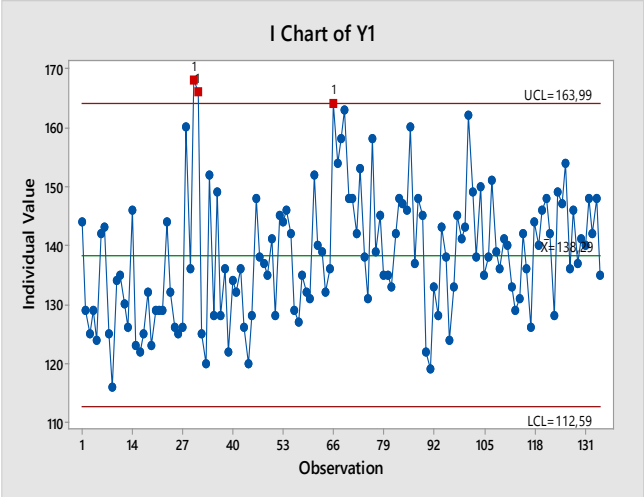

Fig.3 *Individual control chart of strength of tablet (Y_1)*

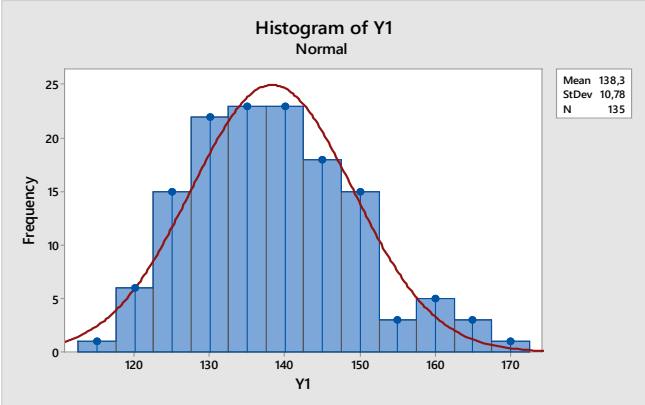

Fig.4 *Histogram of strength of tablet (Y_1)*

3.2. Control chart and histogram for quality parameter Dissolution after 150 minutes (Y_6).

The control chart of drug dissolution after 150 minutes is presented in Fig. 5. There are three point outside the control limits (points 11, 61 and 129). The data are out of statistical control. It is necessary to take measures to eliminate the reasons for this process behavior. The histogram shows the empirical distribution of the data (Fig.6).

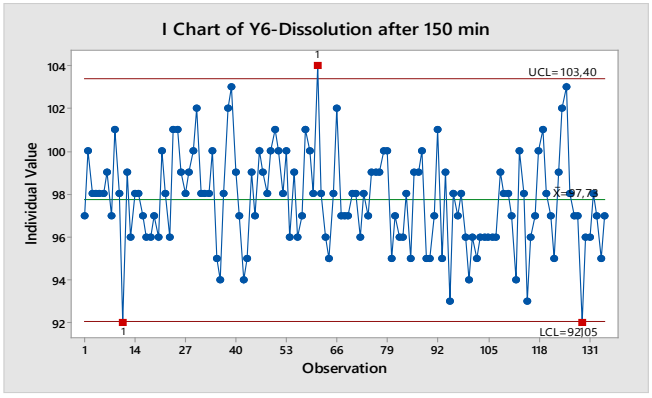

Fig.5 *Individual control chart of the drug dissolution after 150 minutes (Y_6)*

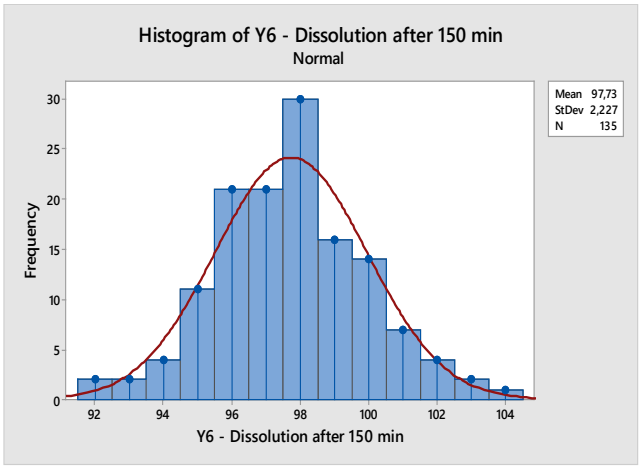

Fig.6 *Histogram of the drug dissolution after 150 minutes (Y_6)*

3.3. Control chart and histogram for quality parameter Dissolution after 165 minutes (Y_7).

The control chart of the drug dissolution after 165 minutes is presented in Fig. 5. There are three point outside the control limits (points 11, 61 and 129). The data are out of statistical control. It is necessary to take measures to eliminate the reasons for this process behavior. The histogram shows the empirical distribution of the data in Fig. 8.

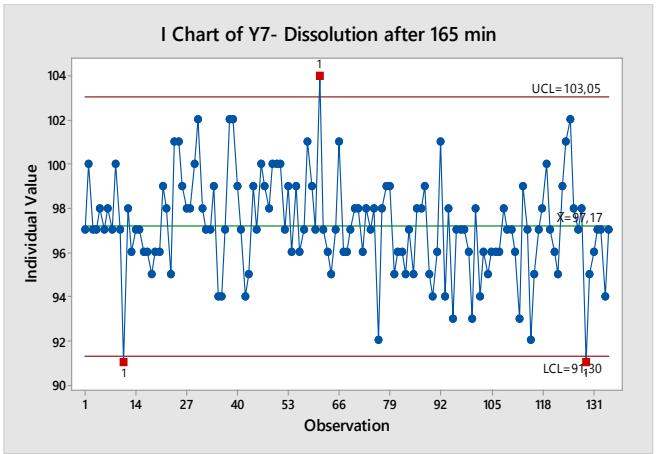

Fig.7 *Individual control chart of the drug dissolution after 165 minutes (Y_7)*

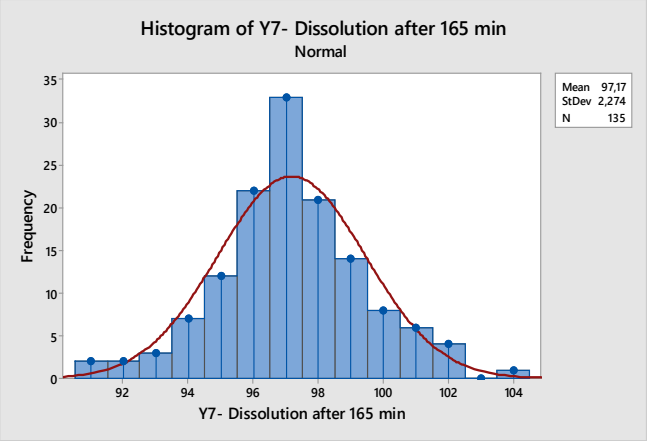

Fig.8 *Histogram of the drug dissolution after 165 minutes (Y_7)*

4. Conclusions

In the present work experimental data from 135 batches are obtained. The studied quality parameters are strength of tablet (Y_1), assay of the active substance (Y_2), acid resistance (Y_3), disintegration (Y_4), water content (Y_5), dissolution after 150 minutes (Y_6) and dissolution after 165 minutes (Y_7). The data are processed with statistical analysis. By applying factor analysis, three significant latent factors for the variables' variation were defined. The control charts and histograms show that the quality parameters – strength of the tablet, drug dissolution after 150 minutes and drug dissolution after 165 minutes, are not under statistical control, because there are points outside the control limits. It is necessary to take actions for the elimination of the special causes for such variation.

References

[1]. Eudralex Volume 4 EU Guidelines to Good Manufacturing Practice Medicinal Products for Human and Veterinary Use, Chapter 6: Quality Control

[2]. Кирков А., Организация и икономика на фармацевтичното производство, София 1981г.

[3]. Myers R. H., D. C. Montgomery, G. G. Vining, C. M. Borror, S. M. Kowalski. Response Surface Methodology: A Retrospective and Literature Survey. Journal of Quality Technology, 36,1, 2004, pp. 53-77.

[4]. Vuchkov I. N, Boyadjieva L. N. Quality improvement with design of experiments, The Netherlands, Kluwer Academic Publishers, 2001.

[5]. Димитрова З., История на фармацията, София 2003

[6]. Osborne J. W. What is Rotating in Exploratory Factor Analysis?Practical Assessment, Research & Evaluation, Vol 20, No 2, 2015

[7]. Christopher Burgess, Standard Operating Procedure – Laboratory Data Management Guidance