Open Journal of Statistics, 2012, 2, 269-273
http://dx.doi.org/10.4236/ojs.2012.23032 Published Online July 2012 (http://www.SciRP.org/journal/ojs)

Sample Size Determination and Statistical Hypothesis Testing for Core Centration in Press Coated Tablets

Pierre Lafaye de Micheaux[1], Vincent Lemaire[2]
[1]Département de Mathématiques et Statistique, Université de Montréal, Montréal, Canada
[2]Pfizer Inc., San Francisco, USA
Email: lafaye@dms.umontreal.ca, vincent514@gmail.com

Received April 11, 2012; revised May 15, 2012; accepted May 30, 2012

ABSTRACT

A novel statistical approach to evaluate the manufacturing quality of press coated tablets in terms of the centering of their core is presented. We also provide a formula to determine the necessary sample size. This approach is applied to real data.

Keywords: Core Centration; Statistical Hypothesis Testing; Dry-Coated Tablets; Sample Size

1. Introduction

Dry-coated tablets provide an adequate and inexpensive technology for the development of controlled-release tablets. The combinations of different compositions between the core and the coat allows for a large variety of design of the release profile [1,2]. It is important that the core is well centered inside of the tablet to make the tablet more light-stable, more water-stable or to preserve the release properties [2-4]. The positioning of the core varies slightly from tablet to tablet during the manufacture process [5-9]. A quality assessment process needs then to be established to guarantee that the centering of the core stays within certain specifications. Current methods used for evaluating core centration are either imprecise, where the core position is measured "by hand" [1,10], or expensive, where costly diagnostic equipment is required [3]. An option to obtain well-centered cores is to use one-step dry-coated tablets [3,11] but the technology is rather new and not implemented everywhere. In this paper, we propose a relatively inexpensive and reliable methodology to monitor core centration during the manufacture of dry-coated tablets. The methodology is based on the cutting of a small sample of tablets (see Section 2), dyeing of the cut surface to distinguish the core from the coat, taking pictures of the cut surface with a digital camera, analysis of the pictures using image processing algorithms to extract core positioning parameters (see Section 3), and a statistical analysis of these data to provide core positioning information for the whole production batch and to compute the minimal sample size required to give a meaningful statistical answer. The goal of this paper is to present in details the statistical method that was developed (see Section 4) and an example of application to real data (see Section 5).

2. The Core Centration Problem

The notations for the dimensions of the diameter and thickness of the tablet and of the core are given on **Figure 1**.

Generally, the exclusion criterion for deciding if the centering of a core is acceptable or not is if the minimum distance m between the edge of the core and the border of the tablet is smaller to some fixed distance m_0, let's say $1/10^{\text{th}}$ of $\min\{(D-d)/2,(E-e)/2\}$.

We will assume that, for practical reasons (e.g. to prevent a breaking of the tablet), it is only possible to make one cut per tablet. The two types of cuts useful to measure the quality of core centration and easy to perform are the longitudinal cut and the transverse cut, as schematized on **Figure 2**.

Once several tablets have been cut, transversely or longitudinally, we measured various displacement quantities, as illustrated on **Figure 3**, using a *modus operandi* described in the next section. The needed measures of position and distance are:

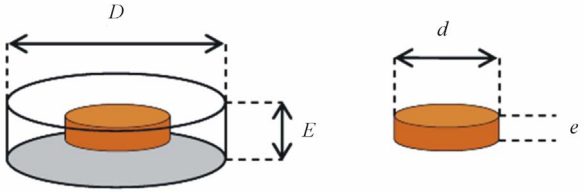

Figure 1. Diameter and thickness for the tablet and core.

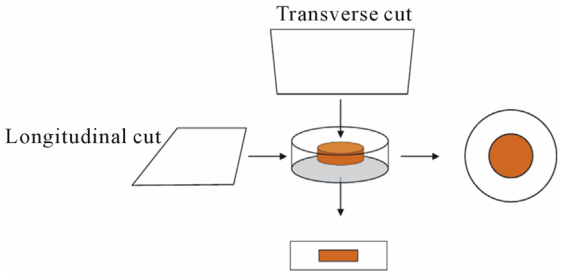

Figure 2. Schematization of longitudinal and transverse cut.

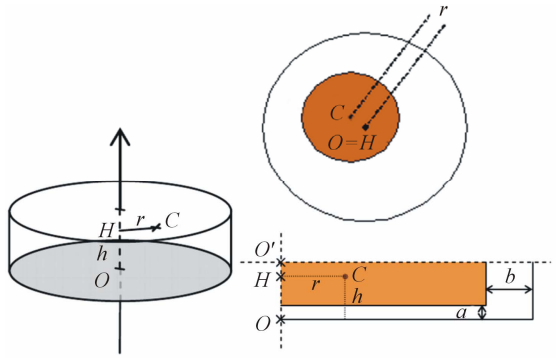

Figure 3. Parameters of interest to determine core off-center. For symmetry reasons, the left and upper parts of the transverse cut may be omitted.

- O is the center of the tablet on the bottom surface of it, while O' is the real center of the tablet;
- C is the position of the center of the core in the tablet and H is its orthogonal projection on the rotational axis of symmetry of the tablet;
- h is the distance (measured on the transverse cut) from H to the bottom surface of the tablet (along the axis of rotation), *i.e.* distance OH;
- r is the distance (measured on the longitudinal cut) from the center of the core C to the rotational axis of symmetry of the tablet;
- The distances $a := a(h) = h - e/2$ and $b := b(r) = (D-d)/2 - r$ that will be used in Section 4 to build the statistical test of core centration.

Moreover, note that we will suppose that vertical and horizontal displacements are independent, and we will also neglect tilt movements of the core, which is a realistic assumption as can be seen on **Figure 5**.

3. A Pattern Recognition Tool

The cut tablets are placed on a tray and two pictures are taken, one for the longitudinal cut (see **Figure 4**) and one for the transverse cut (see **Figure 5**). After dyeing, the core appears dark brown and the outer layer appears beige. Using Matlab and its Image Processing Toolbox, we developed algorithms for image recognition to identify the border of each tablet and each core. The necessary positioning parameters are then computed and used for the statistical analysis.

Figure 4. Picture of longitudinally cut tablets.

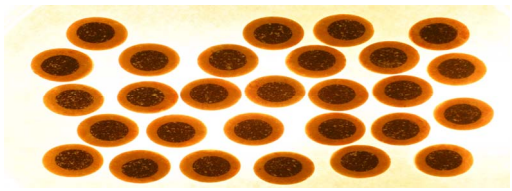

Figure 5. Picture of transversely cut tablets.

4. A Statistical Hypothesis Test

4.1. The Statistical Formulation of the Question

Let M be a random variable giving, for each observed tablet, the minimum distance m between the edge of its core and the border of this tablet. Let m_0 be a reference distance below which a tablet is declared unacceptable and let $p_{m_0} = \mathbf{P}[\mathrm{M} > m_0]$ be the unknown probability that M is above the m_0 threshold.

The objective is to significantly prove, namely with a small and controlled error risk (less than some fixed threshold α), that (in the population) $p_{m0} > c = 1 - \varepsilon$ for a fixed and small given tolerance value ε; that is to say that our tablets are correctly manufactured with high probability.

We can formalize this problem in the statistical hypothesis jargon, using the so-called null and alternative hypotheses:

$$H_0 : p_{m_0} \le c \text{ versus } H_1 : p_{m_0} > c. \qquad (1)$$

Following the notations given in Section 2, we define $a(\mathrm{H}) = \mathrm{H} - \frac{e}{2}$ and $b(\mathrm{R}) = \frac{D-d}{2} - \mathrm{R}$, where we will suppose that the random variables H and R are independent. It is easy to see that $\mathrm{M} = \min\left(a(\mathrm{H}), b(\mathrm{R})\right)$.

Since H and R are supposed to be independent, we have

$$\begin{aligned}
\mathbf{P}[\mathrm{M} > x] &= \mathbf{P}\left[\min\left(a(\mathrm{H}), b(\mathrm{R})\right) > x\right] \\
&= \mathbf{P}\left[a(\mathrm{H}) > x; b(\mathrm{R}) > x\right] \\
&= \mathbf{P}\left[a(\mathrm{H}) > x\right]\ \mathbf{P}\left[b(\mathrm{R}) > x\right] \\
&= \mathbf{P}\left[\mathrm{H} - e/2 > x\right]\ \mathbf{P}\left[(D-d)/2 - \mathrm{R} > x\right] \\
&= \mathbf{P}\left[\mathrm{H} > x + e/2\right]\ \mathbf{P}\left[\mathrm{R} < (D-d)/2 - x\right] \\
&= \left[1 - F_{\mathrm{H}}(x + e/2)\right] F_{\mathrm{R}}\left((D-d)/2 - x\right)
\end{aligned}$$

where $F_{\mathrm{H}}(\bullet)=\mathbf{P}[\mathrm{H}\le\bullet]$ and $F_{\mathrm{R}}(\bullet)=\mathbf{P}[\mathrm{R}\le\bullet]$ are the cumulative distribution functions of H and R respectively.

Thus

$$p_{m_0}=\left(1-F_{\mathrm{H}}\left(m_0+e/2\right)\right)F_{\mathrm{R}}\left(\left(D-d\right)/2-m_0\right). \quad (2)$$

4.2. The Statistic Used to Build the Test

Suppose we have a random sample of n_1+n_2 tablets, where n_1 (resp. n_2) tablets have been longitudinally (resp. transversely) cut. So we end up with the random samples $\mathrm{H}_1,\cdots,\mathrm{H}_{n_1}$ and $\mathrm{R}_1,\cdots,\mathrm{R}_{n_2}$, independent copies of H and R respectively. Based on these observations, we can estimate p_{m_0} with the statistic

$$\hat{p}_{m_0}=\left(1-\hat{F}_{\mathrm{H},n_1}\left(m_0+e/2\right)\right)\hat{F}_{\mathrm{R},\mathrm{n}_2}\left(\left(D-d\right)/2-m_0\right) \quad (3)$$

where $\hat{F}_{\mathrm{H},n_1}(\bullet)=(1/n_1)\sum_{i=1}^{n_1}\mathbf{1}(\mathrm{H}_i\le\bullet)$ and $\hat{F}_{\mathrm{R},n_2}(\bullet)=(1/n_2)\sum_{i=1}^{n_2}\mathbf{1}(\mathrm{R}_i\le\bullet)$ are the empirical distribution functions of H and R respectively, with $\mathbf{1}(\bullet)$ being the indicator function, namely $\mathbf{1}(C)$ equals 1 if condition C is true and 0 otherwise.

4.3. Distribution of the Test Statistic

We have $n_1\left(1-\hat{F}_{\mathrm{H},n_1}(x)\right)=n_1-\sum_{i=1}^{n_1}\mathbf{1}(\mathrm{H}_i\le x)$ is a random variable that can take values in $\{0,\cdots,n_1\}$. In fact, one can show that it has a binomial distribution $\mathrm{Bin}(n_1,p_1)$ with

$$n_1p_1=\mathrm{E}\left[n_1-\sum_{i=1}^{n_1}\mathbf{1}(\mathrm{H}_i\le x)\right]=n_1-\sum_{i=1}^{n_1}\mathrm{E}\left[\mathbf{1}(\mathrm{H}_i\le x)\right]$$
$$=n_1-n_1\,\mathbf{P}(\mathrm{H}_1\le x)=n_1\left(1-F_{\mathrm{H}}(x)\right),$$

so $p_1=1-F_{\mathrm{H}}(x)$. Similarly, one can show that $n_2\hat{F}_{\mathrm{R},n_2}(x)$ has a binomial distribution $\mathrm{Bin}(n_2,p_2)$ with $p_2=F_{\mathrm{R}}(x)$.

Thus the test statistic,

$$n_1n_2\hat{p}_{m_0}=n_1\left[1-\hat{F}_{\mathrm{H},n_1}\left(m_0+\frac{e}{2}\right)\right]n_2\hat{F}_{\mathrm{R},n_2}\left(\frac{D-d}{2}-m_0\right)$$

is a random variable taking values in the set $E=\left\{k=ij;(i,j)\in\{0,\cdots,n_1\}\times\{0,\cdots,n_2\}\right\}$, and whose distribution L is the product of two independent binomial distributions $B_1=\mathrm{Bin}(n_1,p_h)$ with $p_h:=1-F_{\mathrm{H}}(m_0+e/2)$ and $B_2=\mathrm{Bin}(n_2,p_r)$ with $p_r:=F_{\mathrm{R}}\left(\left(D-d\right)/2-m_0\right)$. Note that from (2), we have

$$p_{m_0}=p_hp_r. \quad (4)$$

The (unknown) probabilities of distribution L are

$$\pi_k\left(p_{m_0}\right):=\pi_k=\mathbf{P}\left[n_1n_2\hat{p}_{m_0}=k\right],\forall k\in E. \quad (5)$$

Let B_1 and B_2 be two independent random variables with distributions B_1 and B_2 respectively, and let $\Im_k=\left\{(i,j)\in\{0,\cdots,n_1\}\times\{0,\cdots,n_2\};ij=k\right\}$.

We have, for all $k\in E$,

$$\pi_k=\mathbf{P}\left[\mathrm{B}_1\mathrm{B}_2=k\right]=\sum_{(i,j)\in\Im_k}\mathbf{P}\left[\mathrm{B}_1=i;\mathrm{B}_2=j\right]$$
$$=\sum_{(i,j)\in\Im_k}\mathbf{P}\left[\mathrm{B}_1=i\right]\mathbf{P}\left[\mathrm{B}_2=j\right]$$
$$=\sum_{(i,j)\in\Im_k}\binom{n_1}{i}p_h^i\left(1-p_h\right)^{n_1-i}\binom{n_2}{j}p_r^j\left(1-p_r\right)^{n_2-j}$$
$$=\sum_{(i,j)\in\Im_k}\binom{n_1}{i}\binom{n_2}{j}p_h^ip_r^j\left(1-p_h\right)^{n_1-i}\left(1-p_r\right)^{n_2-j}$$
$$=(A)+(B)+(C),$$

where

$$(A)=\sum_{\substack{(i,j)\in\Im_k\\ i<j}}\binom{n_1}{i}\binom{n_2}{j}\left(p_hp_r\right)^i\,p_r^{j-i}\left(1-p_h\right)^{n_1-i}\left(1-p_r\right)^{n_2-j}$$
$$=\sum_{\substack{(i,j)\in\Im_k\\ i<j}}\binom{n_1}{i}\binom{n_2}{j}p_{m_0}^i\left[\frac{p_{m_0}}{p_h}\right]^{j-i}\left(1-p_h\right)^{n_1-i}\left[1-\frac{p_{m_0}}{p_h}\right]^{n_2-j}$$

$$(B)=\sum_{\substack{(i,j)\in\Im_k\\ i>j}}\binom{n_1}{i}\binom{n_2}{j}p_h^{i-j}\left(p_hp_r\right)^j\left(1-p_h\right)^{n_1-i}\left(1-p_r\right)^{n_2-j}$$
$$=\sum_{\substack{(i,j)\in\Im_k\\ i>j}}\binom{n_1}{i}\binom{n_2}{j}p_h^{i-j}p_{m_0}^j\left(1-p_h\right)^{n_1-i}\left[1-\frac{p_{m_0}}{p_h}\right]^{n_2-j}$$

and

$$(C)=\sum_{\substack{(i,j)\in\Im_k\\ i=j}}\binom{n_1}{i}\binom{n_2}{j}\left(p_hp_r\right)^i\frac{\left(1-p_h\right)^{n_1}\left(1-p_r\right)^{n_2}}{\left(\left(1-p_h\right)\left(1-p_r\right)\right)^i}$$
$$=\sum_{\substack{(i,j)\in\Im_k\\ i=j}}\binom{n_1}{i}\binom{n_2}{j}p_{m_0}^i\frac{\left(1-p_h\right)^{n_1}\left(1-\frac{p_{m_0}}{p_h}\right)^{n_2}}{\left(1-\left(p_h+\frac{p_{m_0}}{p_h}\right)+p_{m_0}\right)^i}.$$

4.4. Critical Region and Statistical Decision

Under $\tilde{H}_0:p_{m_0}=c$, we have from (5) $\pi_k=\pi_k(c):=\pi_k^0$. For a pre-specified significance level α (0.05 or 0.1 are classical values), let $\zeta_\alpha(p_h)$ be the largest (positive) value, called the critical value, such that:

$$\mathbf{P}\left[n_1n_2\hat{p}_{m_0}>\zeta_\alpha(p_h)\middle|\tilde{H}_0\text{ is true}\right]=\sum_{\substack{k\in E\\ k>\zeta_\alpha(p_h)}}\pi_k^0\le\alpha. \quad (6)$$

We will reject H_0 (*i.e.* we will show H_1: our tablets are correctly manufactured) with a controlled type I error risk (probability of taking a wrong decision) as soon as

the observed value $n_1 n_2 \hat{p}_{m_0,obs}$ of $n_1 n_2 \hat{p}_{m_0}$ (computed with our observations) will be greater than $\zeta_\alpha(p_h)$. However, since p_h is unknown, we will substitute it with a value $\tilde{p}_h$ given by an expert. This expert likely computed it from a preliminary study in a context where tablets were known (e.g. using more expensive and sophisticated techniques) to be correctly manufactured (*i.e.* $p_{m_0} > c$).

4.5. Determination of Sample Size

We define Δ as a (fixed and known) effect size value greater than c. Under $\tilde{H}_1 : p_{m_0} = \Delta$, we have from (5) $\pi_k = \pi_k(\Delta) := \pi_k^1$. The type II error risk β is defined using the following equation, where $\zeta_\alpha(p_h)$ is the same value as before:

$$\beta = \mathbf{P}\left[n_1 n_2 \hat{p}_{m_0} \le \zeta_\alpha(p_h) \middle| \tilde{H}_1 \text{ is true} \right] = \sum_{\substack{k \in E \\ k \le \zeta_\alpha(p_h)}} \pi_k^1. \quad (7)$$

It is often required that the power $1-\beta$ of the test be at least equal to some fixed value $1-\beta_0$ (e.g. 0.8). Now, using (6) and (7), it is possible to obtain a formula (possibly not in closed form) relating $n_1, n_2, \Delta, c, \alpha, \beta_0$ and p_h.

We will suppose without loss of generality that $n_1 = n_2 = n$. For given values of $\alpha, \beta_0, c, \Delta$ and p_h $(= \tilde{p}_h)$, the necessary sample size to use is thus obtained by solving in the two unknown values $\left(n, \zeta_\alpha(\tilde{p}_h)\right)$ the following system:

$$\begin{cases} \sum_{k \in E; k > \zeta_\alpha(\tilde{p}_h)} \pi_k^0 \le \alpha \\ \sum_{k \in E; k \le \zeta_\alpha(\tilde{p}_h)} \pi_k^1 \le \beta_0 \end{cases}.$$

4.6. Other Approaches to Approximate the Distribution *L*

Two other approaches may be used to approximate the law L. We could resort to bootstrap methods (see [12]) to approximate it by L^* say. Also, under some conditions we can approximate a binomial distribution with a Gaussian distribution. In this case, we could approximate L with $\hat{L}$ the product of the preceding Gaussian distributions whose density is to be determined (knowing the joint density of the two Gaussians).

Here also, we will reject H_0 (*i.e.* we will show H_1: our tablets are correctly manufactured) with a controlled probability of taking a wrong decision as soon as the observed value $n_1 n_2 \hat{p}_{m_0,obs}$ of $n_1 n_2 \hat{p}_{m_0}$ (computed with our observations) will fall too far in the upper low probability regions of L^* or $\hat{L}$.

5. A Real Data Application

We obtained samples of dry-coated tablets. The dimensions of these tablets are given in **Table 1**. We were able to obtain the following h_i and r_i measurements from 58 tablets of Type 1, from which $n_1 = 28$ were cut longitudinally and $n_2 = 30$ were cut transversely:

Table 1. Dimensions of the tablets.

	Type 1	Type 2	Type 3
D	9.53 mm	10.32 mm	12.7 mm
E	5.1 mm	6.53 mm	6.05 mm
d	6.35 mm	7.14 mm	7.94 mm
e	2.8 mm	4.56 mm	5.06 mm

h_i: 2.42413 2.43201 2.50639 2.38244 2.28380 2.40008 2.49024 2.36265 2.35915 2.45703 2.48598 2.42386 2.43107 2.44298 2.41016 2.41928 2.43686 2.29511 2.41397 2.52839 2.37148 2.46547 2.52098 2.46419 2.28514 2.27086 2.52790 2.40166 2.28546 2.39130;

r_i: 0.32041 0.31734 0.22026 0.37408 0.21759 0.52229 0.35651 0.23147 0.33771 0.33981 0.28513 0.25708 0.23791 0.09140 0.26171 0.30145 0.34680 0.25996 0.31784 0.47770 0.06728 0.08803 0.14386 0.26170 0.14298 0.13447 0.19959 0.16734.

We used the **R** software, version 2.11.0 (2010-04-22) [13] and our code is available from the first author. We performed the test using the (reasonable) values $\alpha = 0.05$, $c = 0.73$, $\tilde{p}_h = 0.94$ and $m_0 = 0.89$. We obtained the following results: $n_1 n_2 \hat{p}_{m_0,obs} = 728$ with a *p*-value equal to 0.03647. Thus, it is possible to conclude at level 5% that the tablets are correctly manufactured. Note that choosing the values $\beta_0 = 0.2$ and $\Delta = 0.9$, the sample size needed can be computed as being $2 \times 29 = 58$ tablets.

6. Conclusion

A statistical hypothesis test was developed to evaluate the manufacturing quality of dry-coated tablets, in terms of off-centering of their core. We also presented a formula to compute the sample size needed to get a fixed power. This research could be refined by taking into account the possible tilt movements of the core that have been neglected in this work. Also, sequential analysis and multiple testing problems could be investigated in this context.

7. Acknowledgements

The authors would like to thank the company which provided the samples of tablets.

REFERENCES

[1] P. A. Kulkarni, M. Jaiswal, S. Jawale, S. V. Shirolkar and P. V. Kasture, "Development and Evaluation of Press Coated Tablets for Chronopharmaceutical Drug Delivery

using Gellable and Permeable Polymers," *Der Pharmacia Lettre*, *Scholars Research Library*, Vol. 2, 2010, pp. 482-497.

[2] Y. Ozeki, M. Ando, Y. Watanabe and K. Danjo, "Evaluation of Novel One-Step Dry-Coated Tablets as a Platform for Delayed-Release Tablets," *Journal of Controlled Release*, Vol. 95, No. 1, 2004, pp. 51-60. doi:10.1016/j.jconrel.2003.10.028

[3] M. Ito, T. Suzuki, N. Wakiyama, H. Teramoto, E. Yonemochi and K. Terada, "Development of a Rapid Process Monitoring Method for Dry-Coated Tableting Process by Using Near-Infrared Spectroscopy," *Chemical and Pharmaceutical Bulletin*, Vol. 59, No. 7, 2011, pp. 868-873. doi:10.1248/cpb.59.868

[4] J. Windheuser and J. Cooper, "The Pharmaceutics of Coating Tablets by Compression," *Journal of the American Pharmaceutical Association*, Vol. 45, No. 8, 1956, pp. 542-545. doi:10.1002/jps.3030450812

[5] L. Lachman, P. P. Speiser and H. D. Sylwestrowicz, "Compressed Coated Tablets. I. Measurement and Factors Influencing Core Centration," *Journal of Pharmaceutical Science*, Vol. 52, No. 4, 1963, pp. 379-390. doi:10.1002/jps.2600520415

[6] Y. Ozeki, Y. Watanabe, S. Inoue and K. Danjo, "Comparison of the Compression Characteristics between New One-Step Dry-Coated Tablets (OSDRC) and Dry-Coated Tablets (DC)," *International Journal of Pharmaceutics*, Vol. 259, No. 1-2, 2003, pp. 69-77. doi:10.1016/S0378-5173(03)00208-4

[7] Y. Ozeki, Y. Watanabe, S. Inoue and K. Danjo, "Evaluation of the Compression Characteristics and Physical Properties of the Newly Invented One-Step Dry-Coated Tablets," *International Journal of Pharmaceutics*, Vol. 267, No. 1-2, 2003, pp. 69-78. doi:10.1016/j.ijpharm.2003.07.011

[8] Y. Ozeki, Y. Watanabe, H. Okamoto and K. Danjo, "Development of Dividable One-Step Dry-Coated Tablets (Dividable-OSDRC) and Their Evaluation as a New Platform for Controlled Drug Release," *Pharmaceutical Research*, Vol. 21, No. 7, 2004, pp. 1177-1183. doi:10.1023/B:PHAM.0000033004.88953.bc

[9] S. Rujivipat, "Novel Formulation and Processing Aspects for Compression-Coated Tablets and for the Compression of Polymercoated Multiparticulates," Ph.D. Thesis, Free University of Berlin, Berlin, 2010.

[10] R. Srinivas, H. G. DeKay and G. S. Banker, "New Approach to Development and Manufacture of Enteric Compression Coatings," *Journal of Pharmaceutical Sciences* Vol. 55, No. 3, 1966, pp. 335-340. doi:10.1002/jps.2600550314

[11] S. A. Yehi, A. H. Elshafeey, I. Sayed and A. H. Shehata, "Optimization of Budesonide Compression-Coated Tablets for Colonic Delivery," *AAPS PharmSciTech*, Vol. 10, No. 1, 2009, pp. 147-157.

[12] B. Efron and R. J. Tibshirani, "An Introduction to the Bootstrap, Ser. Monographs on Statistics and Applied Probability," Chapman and Hall, New York, 1993.

[13] Development Core Team, A Language and Environment for Statistical Computing, Foundation for Statistical Computing, Vienna, 2010. http://www.R-project.org

