ORIGINAL ARTICLE

Evaluation of Random Blinded Re-Checking of AFB Slides under Revised National Tuberculosis Control Programme in Solapur District

*Swapnil Vishnu Lale**

District Tuberculosis Officer, Civil Hospital, Solapur - 413002, (Maharashtra), India

Abstract:

Background: One of the important components of revised national tuberculosis control programme is 'Good quality diagnosis, primarily by sputum smear microscopy'. All efforts are made to ensure that the designated microscopy centers function at optimal level. The process of 'Random Blinded Re-Checking' (RBRC) of Acid Fast Bacillus slides is built in the programme. *Objectives*: To study the relationship of different types of errors detected in RBRC with respect to time, place and cost. To study the stability and capability of the process of RBRC. *Methods*: Analysis of secondary data of external quality assessment of Solapur district since January 2006 is supplemented by direct implementation of the programme since April 2011 till date. Data analysis is done using statistical software Minitab version 16. *Results:* Since January 2006 to May 2012; 42191 slides were re-checked in 77 RBRC sessions at District Tuberculosis Center, Solapur. Different types of 69 errors were detected. On-site evaluation and panel testing did not show any discordance. Barshi and Mangalwedha Tuberculosis Units (TU) showed significantly higher number of errors as compared to Karmala TU. ($P<0.002$) Weighted Pareto Chart revealed that the costliest form of errors is high false negatives and low false negatives. *Conclusion*: Detection of errors in RBRC sessions follows Poisson distribution. The process of RBRC is found to be in control and capable of achieving the desired target of detection of errors.

Key words:

DPU- Defects per unit, HFN- High false negative, RBRC- Random blinded re-checking

Introduction:

Revised national tuberculosis control programme (RNTCP) was pilot tested from 1993 to 1997. RNTCP was scaled up in a phase wise manner to cover the entire country by March 2006. It was implemented in Solapur district from April 2002. Directly observed treatment with short course chemotherapy (DOTS) is a core strategy of RNTCP. One of the important components of DOTS strategy is 'Good quality diagnosis, primarily by sputum smear microscopy.' Microscopy also facilitates categorization of patients; monitor the response to the treatment. Quality assured smear microscopy laboratories are established for this purpose. An effective quality assurance system is built in the RNTCP programme. The former consists of 1) Internal quality control, 2) External quality assessment (EQA) and 3) Continuous efforts for quality improvement of laboratory services [1]. The system also provides credibility of laboratory results and motivation of the staff for further improvement in the efficiency. This paper studies the external

quality assessment aspect only.

This paper attempts to describe the distribution of different types of errors detected in EQA programme of Solapur district with respect to time and place. Identification of poor performing tuberculosis unit (TU) helps the programme manager to pay more attention towards that particular unit. Identification of low performance time period may point towards certain determinants. The secondary data available at the district tuberculosis center (DTC) from January 2006 till date is subjected to analysis by *Minitab* software and salient findings are presented in this paper. The main objectives are to study the relationship of different types of errors detected in random blinded re-checking (RBRC) of AFB slides with respect to time, place and cost and to suggest suitable recommendations. To study the stability and capability of the process of RBRC.

Material and Methods:

The population of Solapur district is 43, 15,527. Solapur city has a population of 951118; while the population of Solapur rural is 3364409 [2]. The latter is divided into seven Tuberculosis Units (TU). They are DTC Solapur, Akkalkot, Barshi, Pandharpur, Akluj, Mangalwedha, and Karmala. A total of 33 Designated Microscopy centers (DMC) are functioning in these 7 TUs. One senior tuberculosis treatment supervisor (STS) and one senior tuberculosis laboratory supervisor (STLS) works in each TU. District Tuberculosis Centre (DTC) is responsible for quality assurance of smear microscopy for all these 33 DMCs. In this capacity, DTC performs External quality Assurance (EQA). EQA has three components. On site evaluation (OSE), b. Panel testing, c. Random blinded re-checking of routine slides.

a. **On-site evaluation**- On-site evaluation is conducted at least once a month by STLS at the DMC. The visit includes a comprehensive assessment of laboratory safety procedures, condition of equipment, adequacy of supplies as well as technical components of AFB smear microscopy like preparation, staining and reading of smears.
b. **Panel testing**- This method evaluates individual performance in staining and reading and not all the laboratory activities. Panel testing is conducted by State tuberculosis training and demonstration centre (STDC), Pune at DTC, Solapur.
c. **Random Blinded Re-Checking (RBRC) of routine slides** - This EQA method provides reliable assurance that a district has an efficient AFB microscopy laboratory network supporting RNTCP. Blinded rechecking is a process of re-reading a statistically valid sample of slides from a laboratory to assess whether that laboratory has an acceptable level of performance. Sample is collected using Lot Quality Assurance Sampling (LQAS) method. LQAS requires smaller samples because it doesn't attempt to construct a precise estimate of population parameters. RBRC is conducted every month [3].

The STLS selects from the laboratory register the sample slides for RBRC using LQAS method. The laboratory technician (LT) prepares a list of the slide numbers that are selected by STLS along with results and encloses in them a sealed envelope. LT arranges the

slides in a separate box supplied by District Tuberculosis Officer (DTO) and marks on the top of the box as well as an envelope with the title: LQAS slides, Name of DMC, TU and the month and year. STLS picks up the box and the envelope and hands them over to DTO. DTO codes and interchanges slide boxes among STLS, retaining the sealed envelope in his possession. Thus blinding of slides is ensured by the DTO. STLS reads and records the results for slides. Umpire reading will be done by another STLS selected by the DTO on the basis of merit to resolve the dispute. According to results from the DMC and results from random blinded re-checking; different types of errors emerge. The errors are classified in five categories

a. High false positive (HFP) - Here DMC results shows high bacillary count but actually the slide is negative.
b. High false negative (HFN) - Here the result from DMC is negative, but in re-checking high bacillary count is noted.
c. Low false positive (LFP) - Though DMC has reported scanty bacilli but in re-checking the slide is found to be negative.
d. Low false negative (LFN) - Here the DMC has labeled this slide as negative but during rechecking scanty bacillary count is observed.
e. Quantification error (QE) - Here there is discordance between DMC and re-checking about the quantification. (Scanty and high bacillary count)

DTO gives feedback and corrective action to laboratory technicians through the Medical officer of the DMC. A variable follows a Poisson distribution if the following conditions are met.

a. Data are counts of events. (Data must contain non-negative integers with no upper bound)
b. All events are independent
c. Average rate does not change over the period of interest

The variable 'number of errors' fulfills all these three criteria. Therefore assumption is made that the variable "Total number of errors" follows Poisson distribution during each session of RBRC during entire duration of observation. This distribution is described by one parameter Lambda. This parameter equals mean and variance.

Capability analysis of the process of RBRC

Working definition of Process - A process is a unique combination of instruments, chemicals, methods and technicians involved in producing a measurable output; for example : RBRC programme. All processes have inherent statistical variability which can be evaluated by statistical methods. Process capability is a measurable property of a process to the specification, expressed as process capability index or as a process performance index. The output of this measurement is usually illustrated by a histogram and calculations that predict how many parts will be produced out of specification. Two parts of process capability are

1) Measure the variability of the output of a process (Errors per RBRC session)-
    a. U Chart
    b. Cumulative DPU
2) Compare that variability with a proposed specification (or a target)
    a. Poisson plot
    b. Histogram of distribution of DPU

c. Summary statistic table

This paper analyses the secondary data collected in the DTC since January 2006. Retrospective analysis is supplemented by direct implementation and supervision of the programme since April 2011 till May 2012. While performing the analysis of quality of any process; it is prudent to use popular, established and widely applied quality assessment software. Therefore data analysis is done using statistical software ***Minitab*** version 16.

Results:

Table 1 shows the dates of proficiency panel testing by STDC, Pune. The reports of STLS from all the TUs show acceptable to good performance. The analysis of on-site evaluation check lists from January 2006 to May 2012 did not show any discordance.

Since January 2006 till May 2012, a total of 77 RBRC sessions were conducted in DTC, Solapur. 42,191 slides from 7 TUs and 33 DMCs were re-checked. Different types of 69

Table 1 - Results of OSE and Panel testing

Date of testing	DTC	Akkalkot	Barshi	Pandharpur	Akluj	Mangalwedha	Karmala
01-06-2005	Acceptable	Good	Acceptable	Good	Acceptable	Good	Good
2006	Proficiency panel testing was not conducted during this period.						
2007							
2008							
10-09-2009	Good	Good	Good	Good	Good	Good	Good
13-07-2010	Good	Good	Good	Good	Good	Good	Good
22-08-2011	Good	Good	Good	Good	Good	Good	Good
18-06-2012	Good	Acceptable	Good	Good	Good	Good	Good
Analysis of on site evaluation check lists over the period from 2005 to 2012 did not show any discordance							

Table 2 - Year wise distribution of different types of errors

Year	Total number of slides examined	HFP	HFN	LFP	LFN	QE	Total number of errors
2006	6140	1	6	2	1	4	14
2007	6596	1	5	0	1	9	16
2008	6033	0	0	0	0	8	8
2009	6428	0	0	1	1	6	8
2010	7440	2	2	1	0	3	8
2011	6269	0	5	0	2	3	10
2012	3285	0	3	0	0	2	5
Total	42191	4	21	4	5	35	69
Data upto 31st May 2012							

errors were encountered during the process. Year wise distribution of different types of errors are shown in Table 2. Maximum number of errors were found in years 2006 and 2007. During this period proficiency panel testing of STLS was not conducted by STDC. (Refer to Table 1) Most common type of errors are quantification errors, followed by high false negatives

Before conducting statistical analysis, Minitab offers to graphically explore the data and assess relationships among the variables. Figure number 1 shows the grouped histogram, which displays the histograms for each tuberculosis unit on the same graph. X axis shows the total number of errors and Y axis denotes the density (frequency). Table embedded in the same figure shows TU wise mean number of errors,

Fig. 1 - TU wise distribution of total number of errors

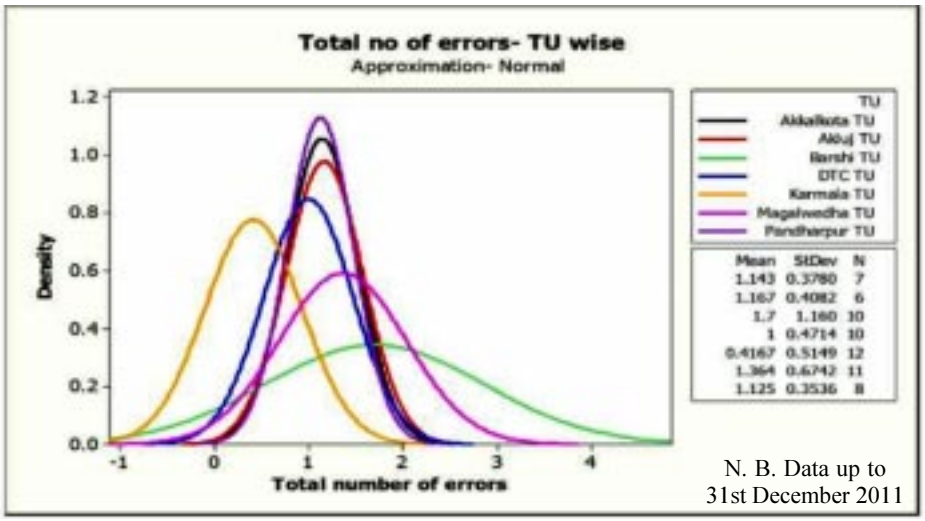
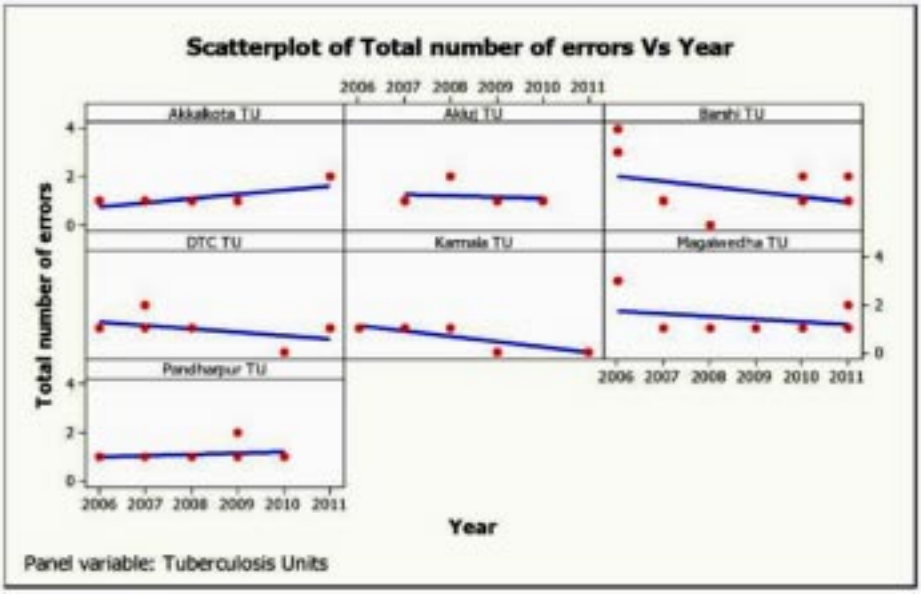


Fig. 2 - Scatter plot of total number of errors vs. years



standard deviation, and total number of errors in each TU. Each TU is depicted in different colour, explanation of which is provided in the legend. This grouped frequency polygon shows that DTC, Akkalkot and Akluj are similar in mean number of errors and spread of errors. In contrast, Barshi TU has more number of errors and more spread out. Later in this paper, efforts are made to detect the statistically significant differences among means using analysis of variance.

This data was obtained from January 2006 to May 2012. But for year wise comparison data up to the completed year 2011 is taken into consideration. As data spans over six calendar years, we suspect whether detection of errors changes over the period of time. To verify this suspicion and eliminate time period as a potentially important factor, the relation between the number of errors over period to years is examined. Fig. 2 shows 7 different panels for each TU. X axis representing the years ranging from 2006 to 2011. By custom, this is a predictor variable. Y axis represents the total number of errors in each year. The blue lines are the regression lines. The points on the scatter plot exhibit no apparent pattern in any of the seven TUs. The regression line for each centre is relatively flat, suggesting that total number of errors are not affected over the years.

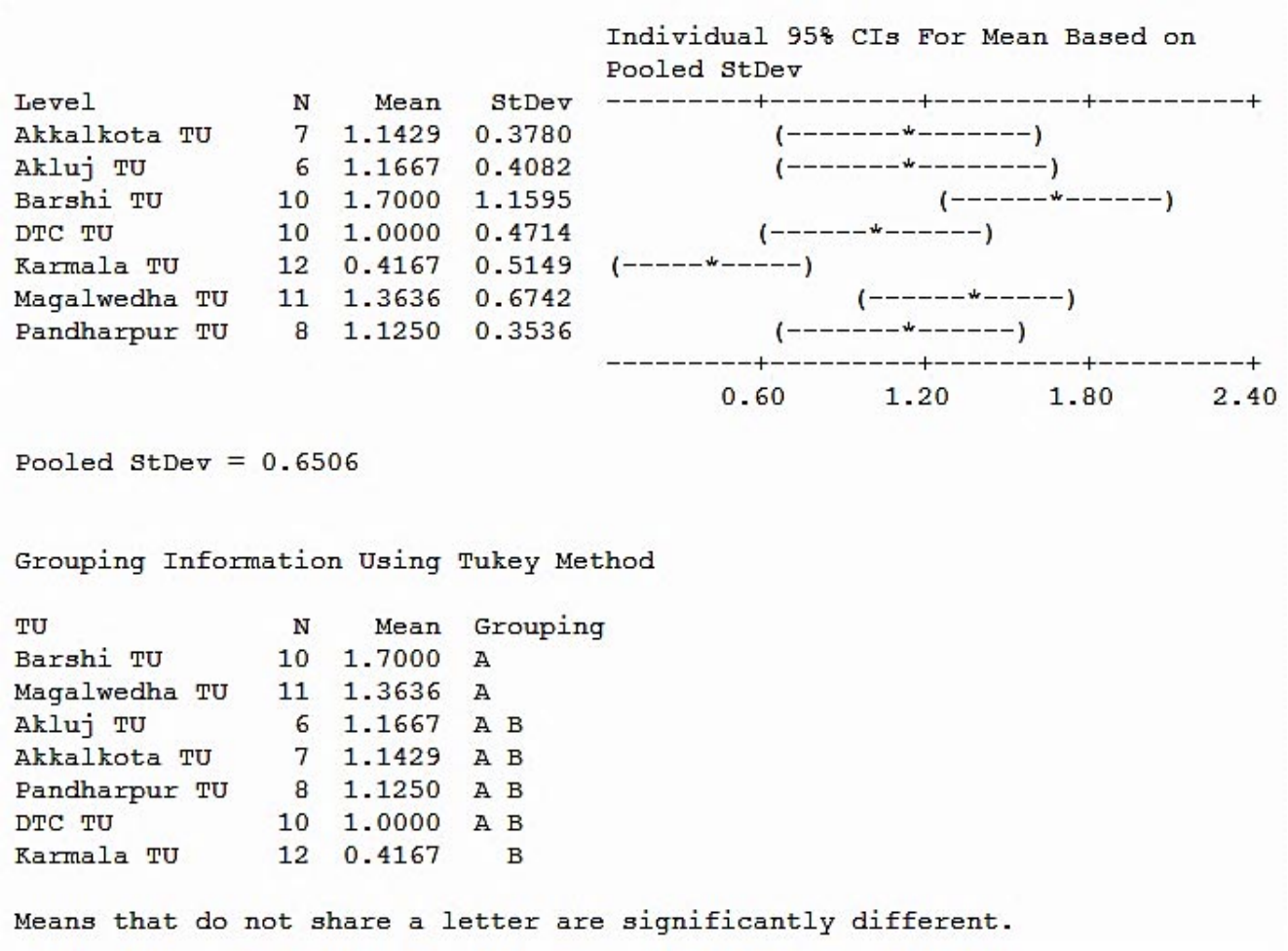
One-way ANOVA: Total no of errors versus TU

Descriptive statistics for each TU such as mean number of errors, standard deviation and total number of errors is provided in table embedded in Fig. 1. Barshi and Mangalwedha TU have more number of errors and higher standard deviations; while Karmala TU has lower mean number of errors and smaller standard deviation. By graphical analysis it seems that the difference in number of errors across different TUs is statistically significant. To verify this, one-way ANOVA test was performed, which tests the equality of two or more means categorized by a single factor. Null hypothesis (H_0) is occurrence of errors is independent of TUs. Table 3 shows that the value of F ratio is 3.98. The calculated value of F is greater than the table value. P value is 0.002. This P value provides sufficient evidence that the mean number of errors is different for at least one TU from the others; when α is 0.05. H_0 is rejected and alternative hypothesis 'Occurrence of errors is dependent on different TUs' is accepted. Tukey's multiple comparison tests is done to test which TU means are different.

Table 3 - One way ANOVA table

Source	Degrees of Freedom	Sum of squares	Mean sum of squares	F Ratio	P
Between the TU	6	10.11	1.68	3.98	0.002
Error (within the TU)	57	24.13	0.42		
Total	63	34.23			
S=0.65, R - Squared = 29.52 %, R-Squared adjusted = 22.10 %, N.B. Data up to 31st December 2011					

Fig. 3- Tukey's multiple comparison tests



Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
Akkalkota TU	7	1.1429	0.3780
Akluj TU	6	1.1667	0.4082
Barshi TU	10	1.7000	1.1595
DTC TU	10	1.0000	0.4714
Karmala TU	12	0.4167	0.5149
Magalwedha TU	11	1.3636	0.6742
Pandharpur TU	8	1.1250	0.3536

0.60 1.20 1.80 2.40

Pooled StDev = 0.6506

Grouping Information Using Tukey Method

TU	N	Mean	Grouping
Barshi TU	10	1.7000	A
Magalwedha TU	11	1.3636	A
Akluj TU	6	1.1667	A B
Akkalkota TU	7	1.1429	A B
Pandharpur TU	8	1.1250	A B
DTC TU	10	1.0000	A B
Karmala TU	12	0.4167	B

Means that do not share a letter are significantly different.

In Fig. 3, Tukey's method is used to create confidence intervals for all pair wise differences between TU level means while controlling the specified family error rate of 5 %. In the 95 % confidence intervals table, intervals of Karmala TU did not overlap the interval of Barshi/ Mangalwedha TU. Also the means of Barshi/ Mangalwedha and Karmala do not share group letters A and B. They are statistically different. Less number of errors are found in Karmala TU while significantly more number of errors are found in Barshi and Mangalwedha TU.

Weighted Pareto chart- Calculating the cost of poor quality

Pareto chart is a special type of bar chart where the plotted values are arranged from largest to smallest. Fig. 4 shows that quantification errors is most common type of error and it is found in 35 instances. It is followed by high false negative (HFN) and low false negative (LFN) observed for 21 and5 times respectively. Least common errors are high false positives (HFP) and low false positives (LFP) found in 4 instances each. Pareto chart is a basic quality control tool used to highlight most frequently

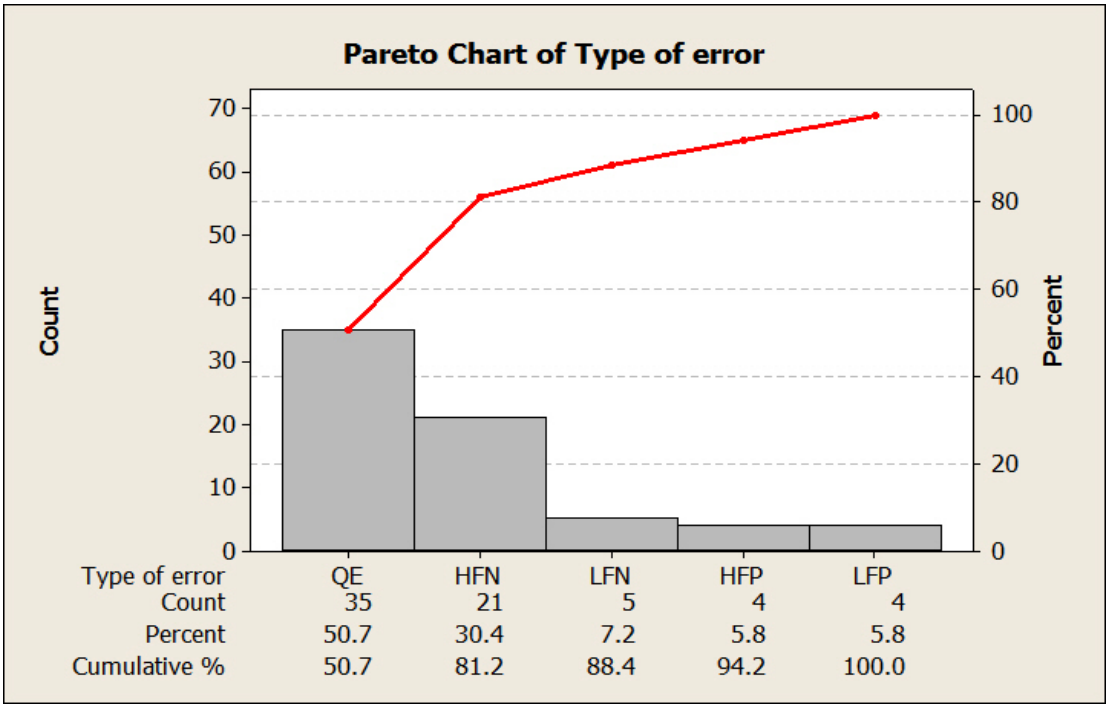
occurring errors. Pareto chart is named after Vilfredo Pareto and his principle of the "80/20 rule". That is 20 % of people contain 80 % of wealth; or 20 % of errors cause 80 % of loss.

A weighted Pareto chart not only considers the frequency of occurrence, but the importance as well. A weighted Pareto chart can account for the loss caused by mislabelling a sputum positive tuberculosis patient as sputum negative or denying the anti-tubercular treatment. It may be the risk of spread of TB or likely death of the patient. In this study data on the frequency of occurrence of error is collected. The cost required for aversion of each DALY due to tuberculosis is $13.3 to $103.5 [4]. So arbitrary valuation in this range for each missed case of tuberculosis is Rs. 4500/-. Same value is used for each case of TB who is unnecessarily treated.

Fig. 4 shows that most costly errors are HFN and LFN. Recommendations should be directed at reducing false negative errors.

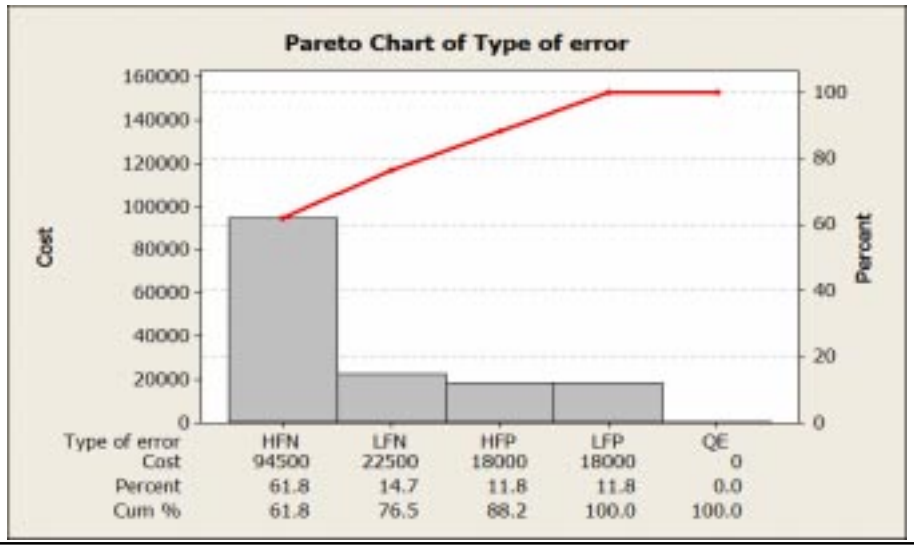
Poisson process - Poisson process describes the number of times an error occurs in finite observation space. Here it describes the number of errors detected in RBRC sessions. Table 4 show the total number of errors detected in monthly RBRC session. 69 errors are spread out in 77 RBRC sessions during a time span of six and half years in purely random manner.

Fig. 4 A - Weighted Pareto chart Showing Frequency



The most frequently occurring error is quantification error (QE). Based just on this information, one may decide to develop an improvement project around reducing QE.

Fig. 4B- Weighted Pareto Chart showing cost of errors:



The most costly defects are HFN and LFN. Based on this more informative data, it is better to develop an improvement project to reduce HFN and LFN type of errors.

N.B. Here data up to 31st May 2012 is taken into consideration.

Table 4 - Confidence Interval for One-Sample Poisson Rate

Variable	Total occurrences	N	Rate of occurrence	95 % confidence interval
Total number of errors	69*	7	0.8961	(0.6972, 1.134)
		7		
Length of observation = 1, * Data up to 31st May 2012				

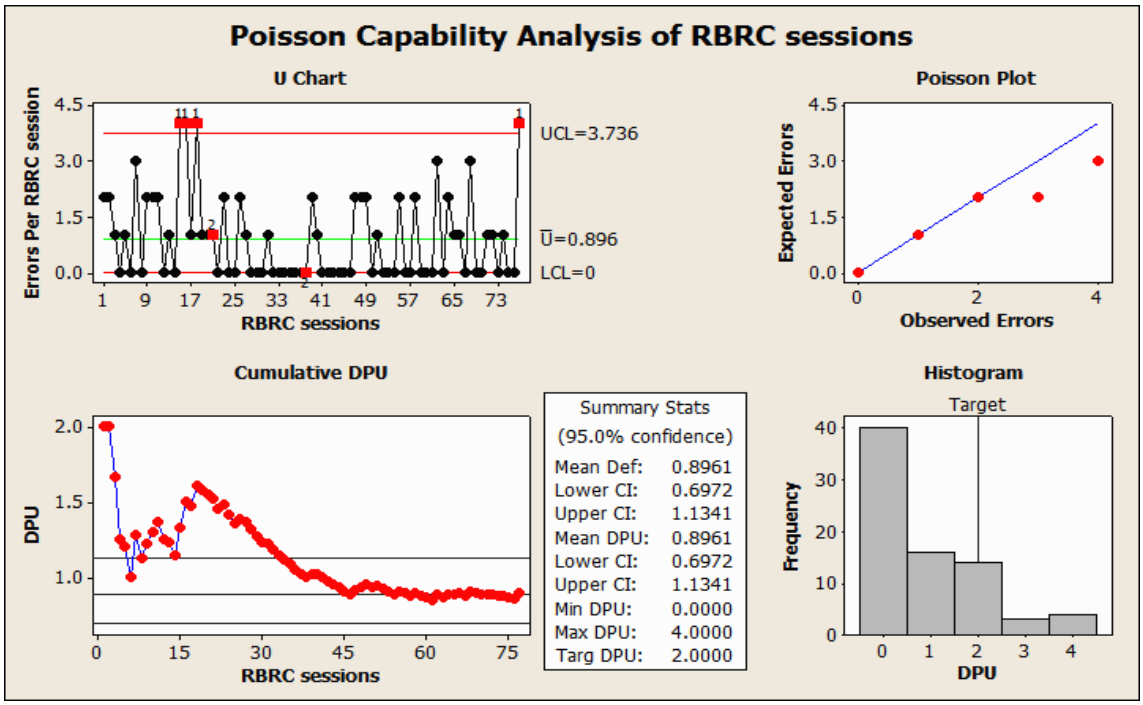
Table 4 shows that in this RBRC process lambda is equal to 0.8961.

Capability analysis of the process of RBRC

Fig. 5 shows five components of capability analysis of process of RBRC.

1) U chart - The U chart is located in the upper left corner of the fig. 5, Process capability analysis. U chart is used to determine whether the number of errors per RBRC session is in control. Plotted points represent the number of errors per session of RBRC. Center line (green) shows the average number of errors per RBRC. Control limits (red) are located 3 δ above and below the center line, and provide a visual means for assessing when the process is out of control. In this study there are 4 points located outside the control limits.
2) Cumulative DPU - The cumulative DPU (defects per unit of measurement) graph helps to determine whether the sample size is adequte to have a stable estimate of the DPU. This graph is located in the lower left corner of the fig. 5. This graph shows that the DPU stabilizes after several samples (at 55th RBRC session onwards) at 0.89. This appears as a flattening of the plotted line. This sta-

Fig. 5 - Poisson capability analysis of RBRC sessions



tistic is central to the capability study, without sufficient data to estimate mean DPU, the analysis can not continue.

3) Poisson plot - Poisson plot is located in the upper right corner of the process capability analysis. Poisson plot plots the expected and observed number of errors. It can be seen from the graph that the plotted points are in the straight line. This proves that the assumption that the data were sampled from a Poisson distribution is correct.
4) Histogram of distribution of DPU - This is located in the lower right corner of fig. 5. The distribution of DPU graph shows that the errors per RBRC follow Poisson distribution. This graph shows a vertical line for the arbitrary target of finding 2 errors per session of RBRC.
5) Summary statistic table - It is located in the lower center of the process capability analysis. It consist of the following
    a. Mean DPU - An estimate of the average number of errors per RBRC (0.8961) as well as confidence intervals for the estimates (0.6972- 1.1341).
    b. Minimum and Maximum DPU - Range is between 0.00 to 4.00
    c. Target DPU - The target number of errors specified by the programme managers. Here the value is 2 errors per RBRC session.

Discussion:

On site evaluation and panel testing failed to detect any error in examination of AFB slides

in Solapur district. These observations differ from those reported by Selvakumar N, et al while studying panel testing in Chennai, South India. They have reported 15 errors out of 360 slides examined in panel testing [5]. This indicates that when the Panel of slides was examined carefully in presence of STDC, Pune team STLSs perform less errors. But Panel testing does not check routine performance. This point towards negligence or casual attitude at routine work and not to technical incompetence. Since January 2006, 42191 slides were rechecked. Different types of 69 errors were detected in RBRC. Similar findings were reported by Malik S et al; while studying RBRC programme in Delhi [6]. Most common type of error has been quantification error followed by high false negatives in the present study. These findings differ from those observed by Yip C.W. et al in a study at Hong Kong. They have reported LFN and LFP as most common forms of errors. This difference may be due to the florescence microscopy procedure they utilized [7] . In a Mexican study; Martinez A et al reported that RBRC detects large number of errors, mostly quantification errors [8]. This study has found out that the distribution of errors is significantly different amongst various tuberculosis units. TU and DMC wise differences are reported in NRHM document of high focus state of Jharkhand [9]. In this study; average DPU has been found to be 0.8961 with a range from 0 to 4 and confidence interval of 0.697 to 1.134. Similar observations have been made by Bais R while using capability index to improve laboratory analytical performance. But they have broadened the range of variability in detection of errors up to 6 δ (standard deviations) on each side of the mean (μ) [10].

Limitations of study:

This study has several limitations. Inherent structure of lot quality assurance sampling does not permit to make precise estimates of the population parameters. So the results of this study may not be generalizable. Certain biases arise because the investigator is a part of the programme management team. For the retrospective part of the study, investigator is unaware of the micro-details of the ground level situation. However he may think of them from his present experience. A precise estimate of the cost of consequences of un-treated TB patients is not available for Indian settings. Therefore Asian values are utilized [4]. They may not be appropriate in local settings. The strengths of this study include the use of various quality assessment tools in ***Minitab*** software. The software is established and widely applied [11]. Using secondary data entails a massive amount of cost and time savings while permitting large sample size.

Conclusions and Recommendations:

More number of errors have been found during 2006-07 period. Incidentally panel testing has not been performed during that period. This finding stresses that panel testing should be conducted regularly. TU wise analysis shows that Barshi and Mangalwedha TU have got more number of errors than Karmala TU. When subjected to statistical analysis, the relationship of TU with number of errors is statistically significant. So programme manager should pay more attention to poorly performing units like Barshi and Mangalwedha TU. Concerned laboratory technician may be trained, infrastructure

of the laboratory may be revamped, and vacant positions need to be filled up.

Pareto chart shows more number of quantification errors. Weighted Pareto chart takes into account the cost of errors. High false negatives and low false negatives turn out to be the costliest forms of errors. As a consequence of false negative error, patients with TB may not be treated, resulting in suffering, spread of TB and death. Intensive phase of treatment may not be extended for the required duration and as a result patient may loose confidence in the programme.

At the time of collection of sputum it is important to make sure that the sample contains sputum, not just saliva, that too in adequate quantity. (at least 2 ml) The Laboratory technician should select thick, purulent particles to make smear and prepare smears correctly- not too thick, too thin or too little material. Slides should be fixed for correct length of time and stained with Carbol fuchsin for the full five minutes. De-colourization with sulphuric acid should not be done too intensively. Smears should be examined for at least five minutes before recording it as negative. They should make sure to label the sputum containers, slides and laboratory forms carefully. Last but not the least cross check the number on the laboratory form and sputum container before recording.

The number of errors detected in each RBRC session seems to follow a Poisson distribution, with rate of occurrence equal to 0.8961. The process of RBRC is in control with fewer number of points located outside the control limits. Cumulative DPU rate stabilizes after 55th RBRC session onwards at about 0.89. The assumption of Poisson distribution is found correct and the target of 2 DPU per RBRC session is feasible. The process of RBRC is capable to achieve desired target of detection of errors.

References:

1. Tuberculosis Control in India17.pdf [Internet]. [cited 2012 Oct 7]. http://tbcindia.nic.in/pdfs/Tuberculosis%20Control%20in%20India17.pdf
2. Solapur District Population Census 2011, Maharashtra literacy sex ratio and density [Internet]. [cited 2012 Oct 7]. http://www.census2011.co.in/census/district/364-Solapur.html
3. Managing the revised national tuberculosis control programme in your area, A training module. First ed. New Delhi: Central TB division, Ministry of health andFamily welfare; 2006.
4. Baltussen R. Cost effectiveness analysis of strategies for tuberculosis control in developing countries. *BMJ* 2005; 331 (7529):1364.
5. Selvakumar N, Sivagamasundari S, Prabhakaran E, Govindaraju R, Perumal M, Granich R, Wares F, Chauhan LS, Narayanan PR. Storage of heat-fixed unstained sputum AFB smears for panel testing in a tuberculosis unit in South India. *Int J Tuberc Lung Dis* 2005; 9(2):223–5.
6. Malik S, Hanif M, Chopra KK, Aggarwal N, Vashist RP. Evaluation of a new quality assessment strategy for blinded rechecking of random sputum smears for TB in Delhi, India. *Southeast Asian J Trop Med Public Health* 2011; 42(2):342–6.
7. Yip CW, Chan MY, Cheung WF, Yu KW, Tang HS, Kam KM. Random blinded rechecking

of sputum acid-fast bacilli smear using fluorescence microscopy: 8 years' experience. *Int J Tuberc Lung Dis* 2012; 16(3):398–401.

8. Martinez A, Balandrano S, Parissi A, Zuniga A, Sanchez M, Ridderhof J, Lipman HB, Madison B.. Evaluation of new external quality assessment guidelines involving random blinded rechecking of acid-fast bacilli smears in a pilot project setting in Mexico. *Int J Tuberc Lung Dis* 2005; 9(3):301–5.
9. Jharkhand_Report .pdf [Internet]. [cited 2012 Oct 7]. http://www.mohfw.nic.in/NRHM/Documents/High_Focus_Reports/Jharkhand_Report%20.pdf
10. Bais R. Use of Capability Index to Improve Laboratory Analytical Performance. *Clin Biochem Rev* 2008; 29(Supplement (i)):S27–S31.
11. Henderson GR. Six Sigma Quality Improvement with Minitab. John Wiley and Sons; 2011.

***Author for Correspondence**: Dr. Swapnil Vishnu Lale, District Tuberculosis Officer, Civil hospital, Solapur - 413002, (Maharashtra), India. Cell No. - 9421958419, E mail- swsipn@gmail.com*