District Level-Based Mean Differences in Recovery Speed from COVID-19 Infection in Mpumalanga Province

Professor Bhekizizwe Mdlovu
E-mail of corresponding author: bhekizm@gmail.com

Abstract
The aim of this analysis was to examine mean differences of the speed of recovery (days) from COVID-19 infection by case-patients based in three different districts in Mpumalanga province. A sample of 5723 case-patients distributed across three districts in the province; namely Gert Sibande, Ehlanzeni and Nkangala. Using the date at which the result confirming positivity for each case was received and the date at which discharge occurred, the DATEDIF() function in Microsoft Excel was used to calculate the speed of recovery, measured by number of days. The speed (time) from infection to recovery was thus measured as the number of days from first positive to first negative SARSCoV-2 PCR test result. Data was processed in Statistical Package for Social Sciences (SPSS) version 21 for windows prior to conducting statistical analysis. The mean differences in the speed of recovery across the three districts were analysed using the mean comparison Analysis of Variance (ANOVA) technique. Mean and standard deviation (mean ± sd) statistics show that cases in Gert Sibande had the quickest average recovery speed of 16.43 ± 7.14 days, followed by Ehlanzeni with 17.07 ± 7.18 days, while cases in Nkangala had the marginally longest recovery time of 17.09 ± 7.56 days at 95% confidence interval. The $F_{(2, 531.96)}$ statistic (= 4.297; $p < 0.05$) and the Tukey HSD post-hoc results indicate significant differences in the speed of recovery by case-patients in the three districts. The mean statistics results demonstrate that the mean recovery speed of case-patients in Gert Sibande differed significantly from average recovery speeds of case-patients in Ehlanzeni and Nkangala.
Keywords: Recovery Speed from COVID-19, Mpumalanga
DOI: 10.7176/JHMN/80-17
Publication date: September 30th 2020

1. Introduction
Human coronaviruses have become common throughout the world. Many different coronaviruses are identified in animals, while merely a small number of them can cause disease in human beings (Department of Health, 2020). The coronaviruses are basically a large family of viruses which may cause respiratory infections ranging from common colds to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). The global pandemic of coronavirus disease 2019 (COVID-19) was first reported on 31 December 2019 by the World Health Organization (WHO) country office following a cluster of pneumonia cases in Wuhan City, Hubei Province of China (National Institute of Communicable Diseases, 2020).

SARS-CoV-2 has been confirmed as the causative virus of COVID-19, which has rapidly become a global pandemic continuing to spread across countries in the world. Most of case-patients initially identified were dealers and vendors at a seafood, poultry and live wildlife market in China (WHO, 2020). In line with the global statistics, people who are most at higher risk are the elderly, individuals with co-morbidities and healthcare workers (Department of Health, 2020). The spread of the disease is believed to happen primarily via respiratory droplets produced when an infected person coughs or sneezes, similar to how influenza and other respiratory pathogens spread. Thus, the majority of cases have occurred in people with close physical contact to cases and healthcare workers proving healthcare to patients with COVID-19 (National Institute of Communicable Diseases, 2020).
Treatment of COVID-19 is mainly supportive in form provision of oxygen for patients with some shortness of breath, coupled with other associated symptoms like fever, dry cough and sore throat, Human antibiotics do not treat viral infections, but such antibiotics may be required if a bacterial secondary infection develops (Lechien, et al., 2019). In pursuit of efforts to curb the rapid spread of the disease, on 1st April 2020, the Ministry of Health in South Africa launched 60 new mobile laboratories to boost the country's capacity to test for COVID-19. Sampling and testing units, procured by the National Health Laboratory Service (NHLS) have been deployed nationwide to all priority districts and metros. The United States (US) President's Emergency Plan for AIDS Relief (PEPFAR)-funded District Support Partners (DSPs) is supporting all the provinces in the programme (Department of Health, 2020).

1.1. Research Objective
- To determine mean differences in the speed of recovery (days) from COVID-19 infection by case-patients based in three different districts in Mpumalanga province.

1.2. Research Question

- Are there mean differences in the speed of recovery (days) from COVID-19 infection by case-patients based in three different districts in Mpumalanga province?

1.3. Research Hypothesis

- There are mean differences in the speed of recovery (days) from COVID-19 infection by case-patients based in three different districts in Mpumalanga province.

1.4. Significance of the Study

The results from this analysis are expected to contribute greatly towards developing mechanisms that can be deemed relevant to help curtail the geographical spread of the disease in the province. From a monitoring standpoint, findings are expected to provide insights on the key areas where mapping and tracking would need to be continuously conducted using dashboards and geographical information systems (GIS).

2. Materials and Methods

2.1. Design

A descriptive research design was applied in the analysis. The categorical factor were the districts in the province, while and dependent variable was recovery speed measured in number of days.

2.2. Population and Sample

The population for the study was patients who got infected and first tested positive from the first SARS CoV-2 PCR test and later tested negative to SARS-CoV-2 PCR test after a certain period. The sample consists of 5 723 clinically confirmed recovered cases as at 21 July 2020.

2.3. Data

The secondary dataset used in the analysis was obtained from Mpumalanga Department of Health (MDoH). Data cleaning and processing was initially conducted in Excel and Statistical Package for Social Sciences (SPSS) prior to conducting statistical analysis. The variable "recovery speed (days)" was calculated by inserting the date at which the result confirming positivity was received and the date discharge occurred into the DATEDIF() function in Excel. Thus, time from infection to recovery is measured as the number of days from first positive to first negative SARSCoV-2 PCR test result.

2.4. Treatment

The sample was categorised into three districts; namely Gert Sibande, Ehlanzeni and Nkangala.

2.5. Statistical Analysis Technique

The univariate one-way Analysis of Variance (ANOVA) statistical technique was applied to examine mean differences in the recovery speed (days) from COVID-19 infection; based on the following function.

$$F\left(z_{ij}, x_{ij}\right) \leq C \quad (1)$$

where F represents the function which transforms x into z; with z denoting recovery speed (days) of the i^{th} case in district j; x represents the i^{th} positivity and discharge record applied to district j; and C denotes a positive scalar, which overall further reduces to:

$$RS_{ij} = \lambda_j + \lambda_j D_{ij} + \varepsilon_{ik}\,; \quad (2)$$

where RS is the recovery speed (days) of the i^{th} case in district j; D represents the district on the i^{th} case in district j; and λ_j represents the average speed of recovery attributed to a particular district.

Descriptive statistics were computed to analyse the means, standard deviation and standard error estimates; while the ANOVA Tukey HSD post hoc test was used to examine whether any significant differences existed between the patients' mean recovery speeds for each of the three districts in the province at 5 percent level of significance.

3. Results and Analysis

This section comprises of two sub sections namely 3.1 which presented descriptive statistics for the study focusing on measures of central tendency and dispersion. In Section 3.2, Analysis of Variance (ANOVA) was discussed in somewhat detail.

3.1. Descriptive Statistics

Table 1: Descriptive Statistics

Recovery speed (days)								
District	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Gert Sibande	1794	16.43	7.139	.169	16.10	16.76	4	88
Ehlanzeni	1258	17.07	7.177	.202	16.68	17.47	1	69
Nkangala	2671	17.09	7.562	.146	16.80	17.38	3	86
Total	5723	16.88	7.353	.097	16.69	17.07	1	88

Table 1 presents descriptive statistics for the speed (days) at which cases which had COVID-19 infection recovered from the disease across the three districts in Mpumalanga province, namely Gert Sibande, Ehlanzeni and Nkangala. From entire sample of five thousand seven hundred and twenty-three (n = 5 723) COVID-19 infected patients, the largest proportion of 47% (n = 2 671) of the cases were in Nkangala, followed by 31% (n = 1 794) of the cases in Gert Sibande, while the least proportion of 22% (n = 1 258) of the cases were in Ehlanzeni.

Based on mean and standard deviation (mean ± sd) statistics, district comparison of the recovery rate from COVID-19 infection statistics show that cases in Gert Sibande had the quickest average recovery speed of 16.43 ± 7.14 days, followed by patients in Ehlanzeni with an average recovery speed of 17.07 ± 7.18 days while patients in Nkangala had the marginally longest recovery rate of 17.09 ± 7.56 days at 95% confidence interval. At provincial level, the recovery speed of patients from the infection stood at 16.88 ± 7.35 days.

3.2. ANOVA Statistics

Table 2: Test of Homogeneity of Variances

Recovery speed (days)			
Levene Statistic	df1	df2	Sig.
2.496	2	5720	.083

The Levene statistics (F-statistic = 2.496; $p < 0.05$), which tests for the assumption of homogeneity of variance show evidence of existence of a significant difference between the two districts' variances. Therefore, the assumption of homogeneity of variance was not satisfied.

Table 3: ANOVA

Recovery speed (days)					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	531.960	2	265.980	4.927	.007
Within Groups	308809.176	5720	53.988		
Total	309341.136	5722			

Table 3 ANOVA results indicate that statistical evidence that mean recovery speed (days) from COVID-19 infection differed significantly among the three districts ($F_{(2, 531)}$ statistic (= 4.927; $p < 0.05$) at 5% level of significance). To detect which of the three districts mean recovery speeds differed significantly from one another; the Tukey HSD post hoc test was conducted (Table 3). In light of the number of comparisons that were made, the Tukey post hoc approach was applied because of its power to control for alpha inflation.

Table 4: Multiple Comparisons

Dependent Variable: Recovery speed (days)						
Tukey HSD						
(I) District	(J) District	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Gert Sibande	Ehlanzeni	-.646*	.270	.044	-1.28	-.01
	Nkangala	-.662*	.224	.009	-1.19	-.14
Ehlanzeni	Gert Sibande	.646*	.270	.044	.01	1.28
	Nkangala	-.016	.251	.998	-.61	.57
Nkangala	Gert Sibande	.662*	.224	.009	.14	1.19
	Ehlanzeni	.016	.251	.998	-.57	.61
*. The mean difference is significant at the 0.05 level.						

The Tukey post hoc tests results show that the mean recovery speed of case-patients in Gert Sibande differed significantly from average recovery speeds of case-patients in Ehlanzeni and Nkangala. No significant differences existed between mean recovery speed of case-patients in Ehlanzeni and those in Nkangala.

Table 5: Homogeneous Subsets

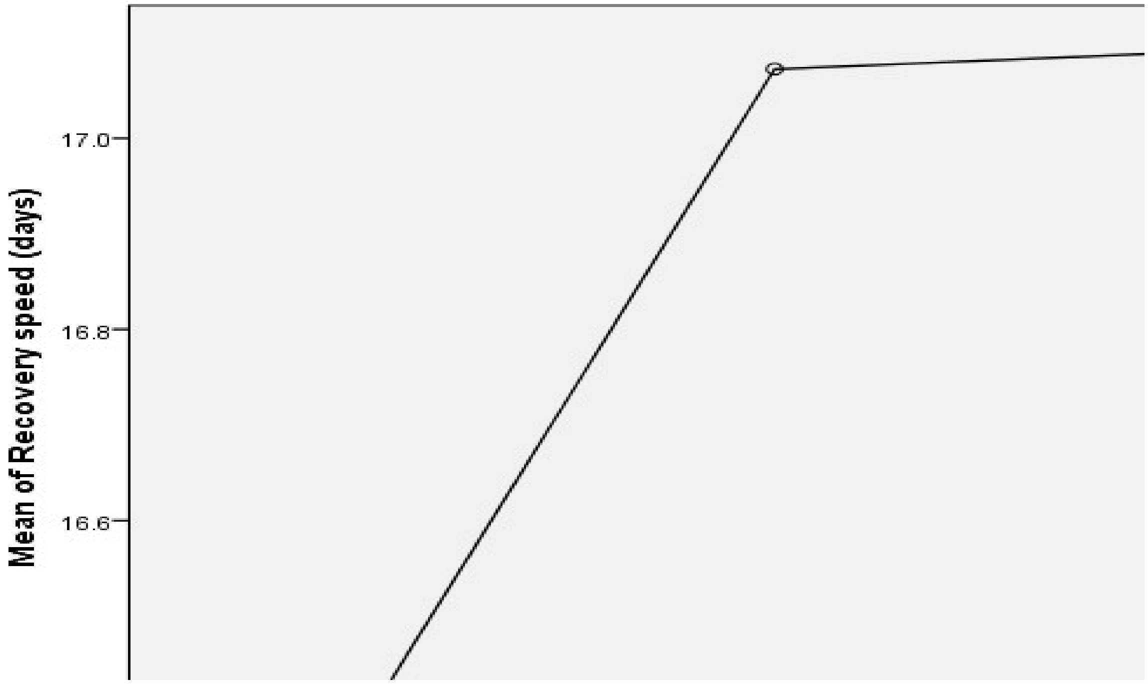
Recovery speed (days)			
Tukey HSD[a,b]			
District	N	Subset for alpha = 0.05	
		1	2
Gert Sibande	1794	16.43	
Ehlanzeni	1258		17.07
Nkangala	2671		17.09
Sig.		1.000	.998
Means for groups in homogeneous subsets are displayed. a. Uses Harmonic Mean Sample Size = 1737.400. b. Group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.			

Homogenous subsets results provide as an alternative approach considered as being more appropriate for computing and displaying the post hoc tests particularly when group sizes are objectively different. The null hypothesis of the subsets establishes that groups computed in the same subset are not significantly different. Results presented in Table 5 indicate significant differences in average speed of recovery case patients in Ehlanzeni and Nkangala districts, while there are no statistically significant differences in the speed of recovery exist among case-patients in Ehlanzeni and Nkangala.

4. Recommendations

Infection prevention and control (IPC) efforts might need to be strengthened in Nkangala where recovery speeds of case-patients in that district are relatively longer than in other districts. The provincial Department of Health should provide assistance to the district support teams (DSTs) aimed at upscaling interventions that prevent the spread of infections during healthcare service delivery in facilities like hospitals, outpatient clinics, dialysis centers and long-term care facilities. Such measures may include rapid identification of suspect cases, rapid screening or triage at initial healthcare facility encounter and rapid implementation of source control, limiting the entry of healthcare workers and visitors with suspected or confirmed COVID-19, cohort patients with suspected or confirmed COVID-19 separately, and safe clinical management.

Appendix: Means Plots



References

Department of Health. (2020). Corona Virus(COVID-19) Pandemic. Republic of South Africa.

World Health Organisation. (2019). Coronavirus disease (COVID-19) pandemic. https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Kai-Wang To, K, Tak-Yin Tsang, O, Chik-Yan Yip, C, Chan, KH, Wu, TC, Man-Chun Chan, J…Yuen, KY. (2020). Consistent detection of 2019 novel coronavirus in saliva. *Clinical Infectious Diseases*. 12 February 2020. ciaa149.
https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa149/5734265external icon

WHO. Clinical management of severe acute respiratory infection when COVID-19 is suspected. 13 March 2020. https://www.who.int/publications-detail/clinical-management-of-severe-acute-respiratory-infection-when-novel-coronavirus-(ncov)-infection-is-suspectedexternal icon

The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. (2020). The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) – China 2020. *CCDCweekly*.10.46234/ccdcw2020.03210.46234/ccdcw2020.032external icon10.46234/ccdcw2020.032external icon 10.46234/ccdcw2020.032external icon

WHO. Infection prevention and control during health care when novel coronavirus (nCoV) infection is suspected. 19 March 2020. https://www.who.int/publications-detail/infection-prevention-and-control-during-health-care-when-novel-coronavirus-(ncov)-infection-is-suspected-20200125external icon

WHO. Rational use of personal protective equipment for coronavirus disease (COVID-19) and considerations during severe shortages. 6 April 2020. https://apps.who.int/iris/bitstream/handle/10665/331695/WHO-2019-nCov-IPC_PPE_use-2020.3-eng.pdfpdf iconexternal icon