Summer Training Report

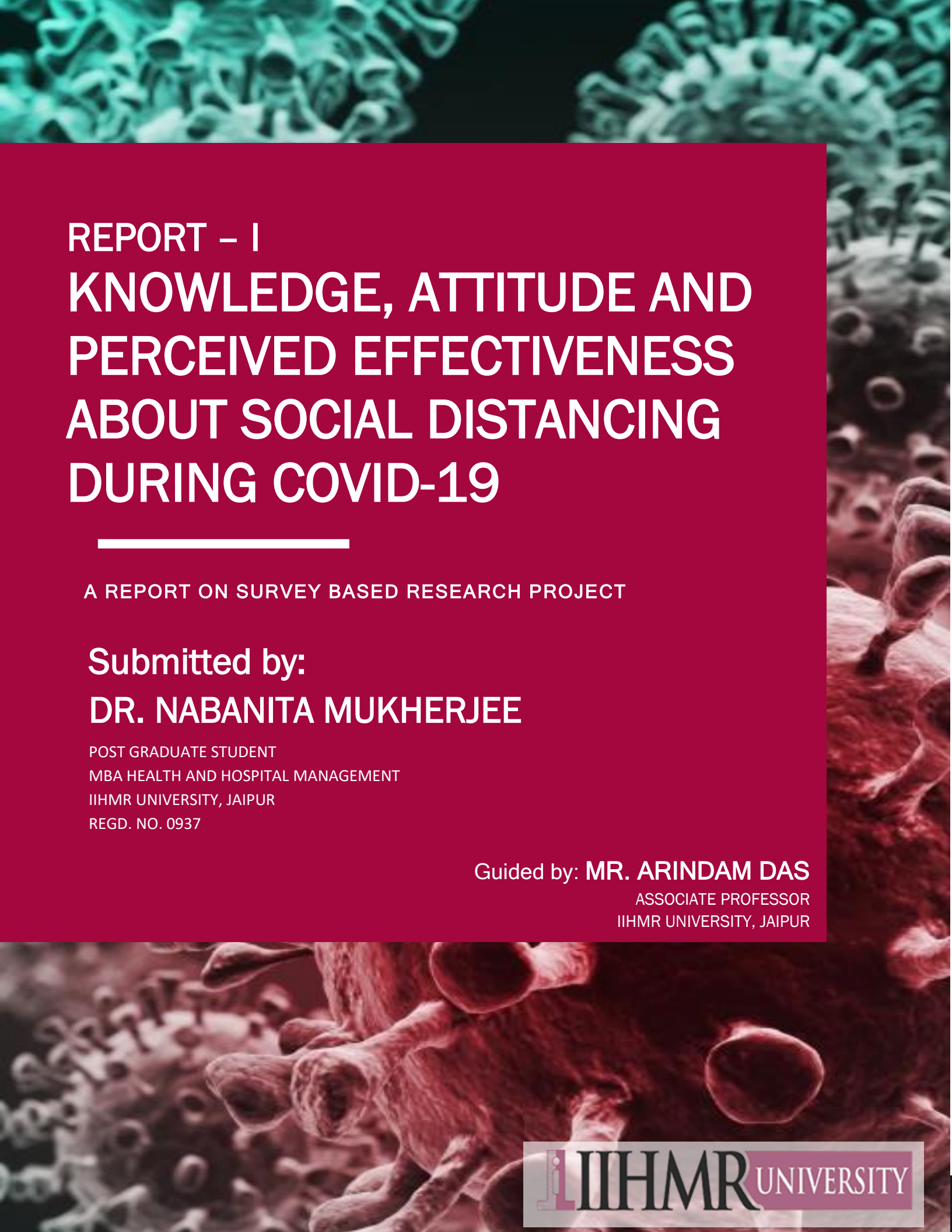
By

Dr. Nabanita Mukherjee

MBA (Health and Hospital Management)

2019-21

IIHMR UNIVERSITY



REPORT – I

KNOWLEDGE, ATTITUDE AND PERCEIVED EFFECTIVENESS ABOUT SOCIAL DISTANCING DURING COVID-19

A REPORT ON SURVEY BASED RESEARCH PROJECT

Submitted by:

DR. NABANITA MUKHERJEE

POST GRADUATE STUDENT
MBA HEALTH AND HOSPITAL MANAGEMENT
IIHMR UNIVERSITY, JAIPUR
REGD. NO. 0937

Guided by: MR. ARINDAM DAS
ASSOCIATE PROFESSOR
IIHMR UNIVERSITY, JAIPUR

IIHMR UNIVERSITY

Acknowledgement

With immense pleasure, I present to you my summer training report part – 1 titled as "**Knowledge, Attitude and Perceived Effectiveness about Social Distancing during COVID-19**" as a part of the 1st year academic curriculum of MBA Health and Hospital management course.

I thank all who in one way or another contributed in the completion of this report.
I would like to extend my profound thanks to my mentor "MR. ARINDAM DAS", under whose expert guidance and critical supervision, I could create and complete this report.
I am also deeply thankful to the respondents of the survey conducted for the research. Their names cannot be disclosed, but I want to acknowledge and appreciate their help and transparency during my research. Their information has helped me complete this report.

Lastly, I thank all those who have directly or indirectly contributed and helped in completion of my report.

DR. NABANITA MUKHERJEE
MBA-HM-24
2019-2021
REGD NO. 0937

TABLE OF CONTENTS

Abstract

Rationale: Social distancing has proven to be the need of hour to break the chain of transmission of the pandemic novel coronavirus. It is essential to assess, to what extent social distancing has been successful in combating the spread of the virus in various countries, especially in India, an overpopulated developing country with extreme cultural variance and huge social interactions. The credibility of managing and holding the infection explosion so far, goes to early preventive public health measures like social distancing, case isolation, quarantine and lockdown being implemented in the country at an early stage. As much as governments across the world and other health authorities including World Health Organization are stressing on the importance of practicing social distancing, it is essential to understand the perception and attitude of the people towards this approach.

Objective: To study the awareness of people to follow social distancing in combating the spread of COVID-19. To study the attitude of the people pertaining to social distancing in controlling the spread of novel coronavirus 2019. To study the barriers in implementing social distancing strategies in an over - populated country like India.

Methodology: Primary data was collected using online questionnaire. An observational study is done. The primary respondents were individuals above 18 years of age who can read and write English having a working internet access with a basic knowledge about healthcare. Convenience Sampling was used. As many as, 413 responses were received. Analysis was done using basic statistical tools like, mean, percentage distribution and standard deviation. Interpretation of primary data collected, was done using Ms Excel and graphical representations.

Results: The results of the survey showed a high level of awareness among the participants about the novel coronavirus infection and the intervention strategies imposed by the Government. A willingness in the attitude to follow social distancing and other preventive strategies, was seen in moderate number of respondents. A moderate to high number of respondents showed some or the other issues that they are facing with regard to their mental health. The barriers in practicing social distancing were identified as well.

Conclusion: People are getting inactive, facing the fear of getting infected, worrying about the economic loss, feeling lagged behind in their lives, getting uncertain about their future plans and many more. There has been numerous instances of violence and stigmatization against the frontline health workers. This scenario is a result of lack of awareness and acceptability of strategies adopted by the Government. Several issues were identified through this study and proposals have been framed to mitigate such issues.

Keywords: Social distancing, COVID-19, awareness, attitude, barriers, mental health, S.W.O.T. Analysis

Introduction

On December 31, 2019, China reported a mysterious pneumonia in the city of Wuhan, to the World Health Organization (WHO) [Error! Reference source not found.]. The disease, as it progressed rapidly to the later part of the world, was then declared as Pandemic on January 30, 2020 by WHO, which then named it as the novel coronavirus 2019 or COVID-19. Since, COVID-19 had symptoms similar to that of the SARS CoV infection, nations across the world, rolled out with precautionary measures similar as that of any contagious respiratory infection. Despite that, the former has been able to roll out over 200 nations globally breeding out 22,85,210 infected cases and claiming 1,55,124 lives, as of 20 April 2020, 2:00am CEST, as shown in Figure.1.[1].

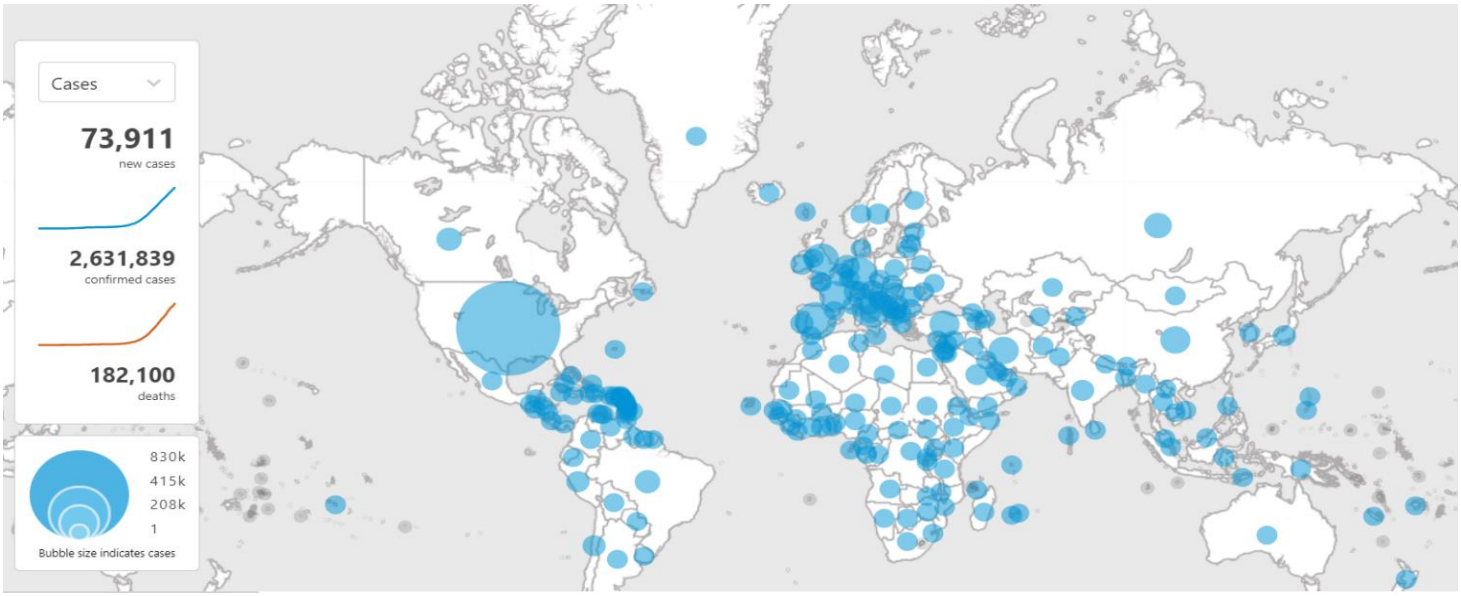


Figure 1: The total number of confirmed cases worldwide with the number of new cases and total number of deaths. Source. WHO [2]

In its inception, the infection caused by the novel coronavirus shows symptoms of a normal fever, fatigue and cough, progressing to a severe respiratory infection in 5-14 days (incubation period) causing difficulty in breathing and fatal pneumonia. Due to these mild symptoms, people might not recognize their condition as “ill” and thus, become a susceptible medium to transmit the infection to others. Younger age people, since being more immune, might surpass the infection stage in a week or two. But similar is not the case for the vulnerable part of the society i.e. older generation and people with pre-existing medical conditions and other debilitating diseases. The virus transmits through droplets when an infected person coughs or exhales within 1 meter of range to another person or on any object or surface followed by touching eyes, nose or mouth by the one who touches that object [Error! Reference source not found.].

Here comes the role of public health measures, like sanitization, wearing masks etc. with the more drastic ones like Social distancing followed by community lockdown. Social distancing has proven to be very effective in controlling the spread of the pandemic COVID-19 [2]. It not only prevents the infection; social distancing also

provides healthcare professionals with the buffer time to treat ample number of patients and ensure that each patient receives proper healthcare. Even though, there is a negative impact on the mental health of people under the effect of social distancing, it is the need of hour to break the chain of transmission of the novel coronavirus. It is essential to assess, to what extent social distancing has been successful in combating the spread of the virus in various countries, especially in India.

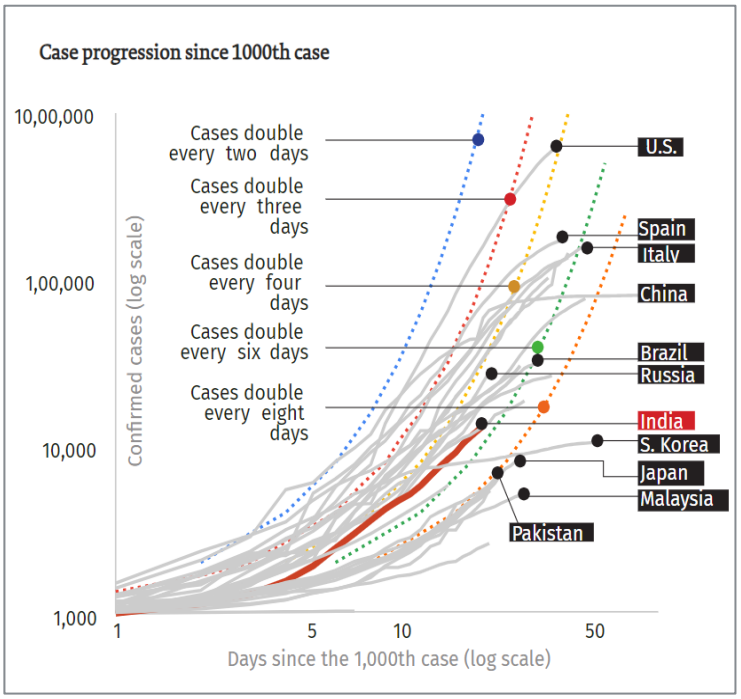


Figure 2: Doubling rate of infection globally since 1000th case. Source. The Hindu [6]

There is a study conducted by Rick Nunes-Vaz [4], where he has visualized the doubling time of new infected cases with the containment measures taken by the country. According to the graphical study done by him, the time taken for the growth of new infection i.e. the doubling time is more in Republic of Korea and Japan, as compared to Italy, USA, UK and Australia. Rick states that the reason behind this outcome are very strict isolation strategies including social distancing and lockdown in Republic of Korea and Japan. Countries like Italy, USA, UK who are among the worst sufferers by the disease, somewhere lacked in implementing or following these social distancing and isolation measures. Another report cited by The Hindu [Error! Reference source not found.], reported the Indian scenario of doubling rate of new infections of COVID-19 as per the research done by Indian Council of Medical Research (ICMR). As shown in Figure.2, the doubling rate of infection in USA is nearly 4 days that is faster in comparison to that of India. The report showed India's doubling rate that there were 987 cases in India on March 28, four days after the declaration of the nationwide lockdown.

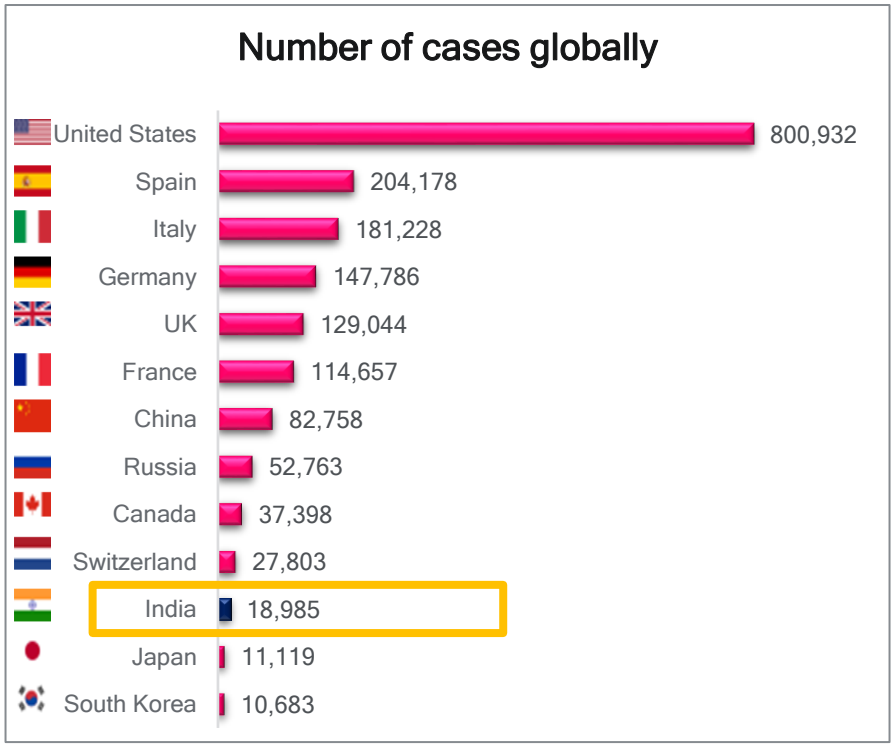


Figure 3: Total number of cases globally as of 22 April, 2020, 2.00 a.m. CEST. Source. Wikipedia [7]

Since then, cases doubled about every four days in the first ten days, between four and five days in the next one week, and between five and six days currently. India is also in a better state with respect to the number of cases in other countries of the world, as per Figure.3. [6]. As of 21 April 2020, 17:00 GMT+5:30, India has 15,122 infected patients and 603 lives has been claimed by the virus [8]. The complete lockdown was imposed by our hon'ble Prime Minister Shri Narendra Modi on March 25, 2020. Since then, several health authorities, media, celebrities etc. have been stressing on the importance of social distancing during this period of lockdown. But there are several gaps in this approach as well. There is a lack of awareness and a lot of barriers in implementing this social distancing strategy in the general population of India. There have been cases of people of hiding in places like Nizamuddin's tabliqi jamaat without proper isolation and precaution, mass gathering of public in Bandra station, lack of proper surveillance in slum areas and other attitudes of people showing careless behavior towards the complete lockdown. Additionally, the mental health of the people is also a matter of great concern and potentially, a barrier against social distancing.

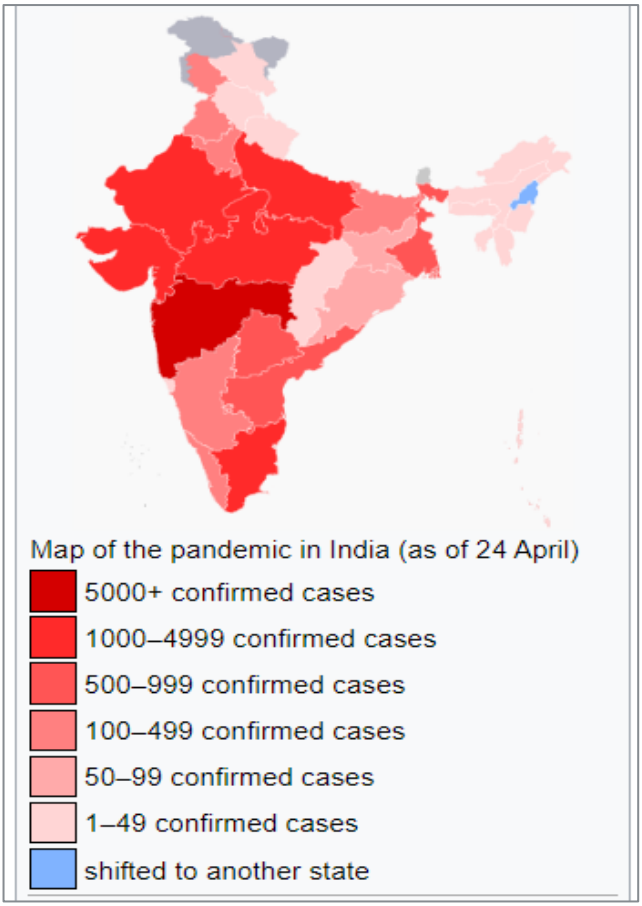


Figure 4: Case wise distribution of confirmed cases of COVID-19 in different states of India as of 21 April 2020, 17:00 GMT+5:30. Source. MoHFW [8]

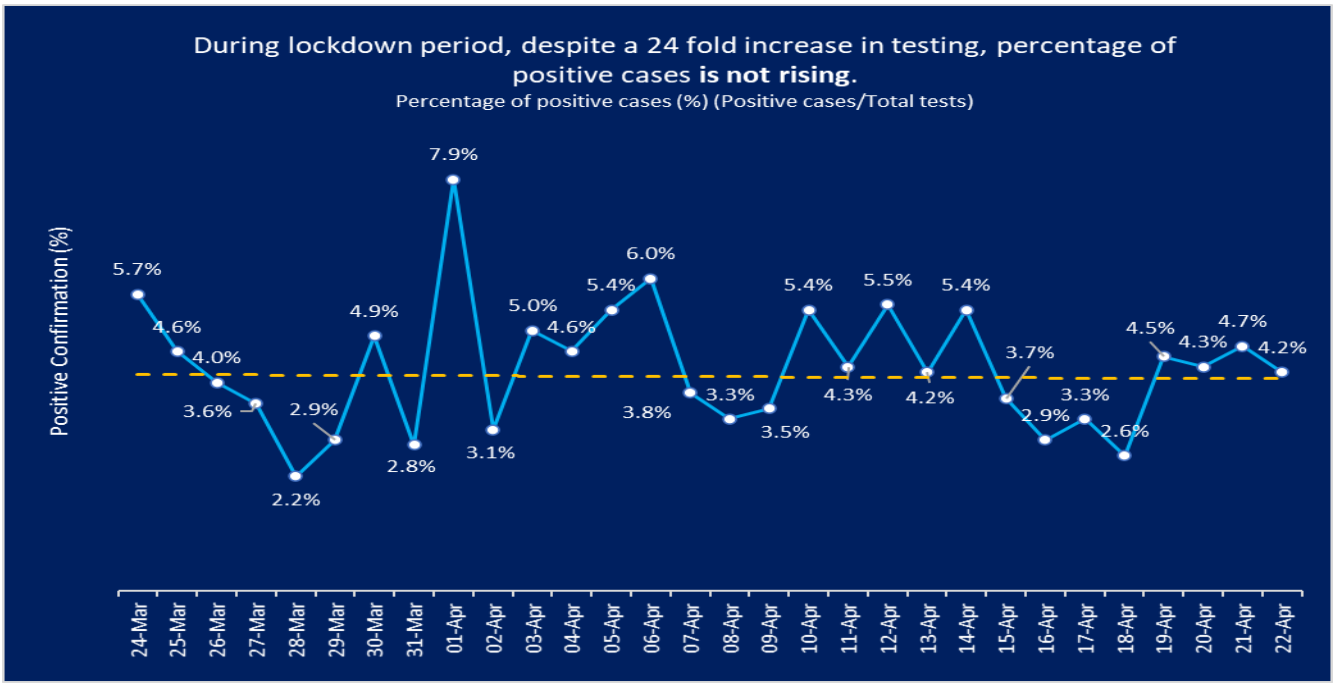
STUDY OBJECTIVE

This study aims to study all aspects of people's perception about the effectiveness of the social distancing measures in India, and the role it has played in reducing the wrath of the pandemic so far. This study aims to achieve the following specific objective;

- Study the awareness of people to follow social distancing in combating the spread of COVID-19.
- Observe the attitude of the people pertaining to social distancing in controlling the spread of novel coronavirus 2019.
- Find out the barriers in implementing social distancing strategies in an over populated and culturally oriented country like India

STUDY BACKGROUND

Social Distancing in Compliance with COVID-19



Graph 1: The percentage of positive cases with respect to the number of tests done in India after the nationwide lockdown was imposed on 23 March, 2020.Source: MoHFW and ICMR. Credit. inshorts [9]

Graph.1 [9] clearly implies that the social distancing measures have positively impacted to reduce the rate of transmission of the COVID-19 infection. Additionally, Indian authors Singh and Adhikari [Error! Reference source not found.] said that the peak of infection would have raised to 0.9 million in 114 days if there were no lockdown measures implemented, since March 25, 2020. They also suggest Four strategies of social distancing and lockdown; First strategy is a complete lockdown and social distancing for a period of 3 weeks or 21 days. Second strategy is to suspend the lockdown for 5 days, then again continue for 28 days. Third strategy includes a consecutive lockdown of 21 days, 28 days and 18 days with a gap of 5 days. Fourth strategy is a 49 days lockdown at a stretch. Number of infections was expected to reduce, and prevent resurgence to a great extent with contact tracing and quarantine methods, with the third and fourth strategy.

In their study [Error! Reference source not found.], Smith and Freedman phrased that Social distancing is effective in places where there is a possibility of substantial community transmission. It, thus, reduces the scope of interaction between individuals of a broad community who are infected but not yet identified and thus, still not isolated from the community. Since, the respiratory infection caused by the novel coronavirus is transmitted through droplets, the proximity of individuals to each other plays a great role in spreading the infection. A study by Joel Koo et.al [10] in their Influenza epidemic simulation model, suggested that a combined intervention, i.e., quarantine, closure of schools and workplace distancing were effective to great extent to reduce the number of infections, in comparison

to the scenario with no intervention. Similarly, the study of Milne and Xie [12] suggests similarity between the public health measures for the novel influenza virus infection and the COVID-19 infection through a community-based simulation model in a city in Australia.

But the authors clarified that in absence of vaccines and treatment for a prolonged time, these measures might attain fatigue. Another study by Kissler et.al [Error! Reference source not found.] outlined the strategies centered on; touch monitoring, quarantine and social distancing, Intensive checking, tracing and isolation of cases; that has allowed transmission control in countries like Singapore and Hong Kong. In an article, Alicia Ault suggested that Multicycle social distancing is an alternative, i.e., instituting social distance before the infection curve bends down to a point where social interactions seem safe to resume [13]. Additionally, Zhang et.al developed SEIQR models that showed early social distancing intervention could dramatically reduce the size of the outbreak, in particular synchronizing the intervention nationally at the earliest possible time [Error! Reference source not found.].

As we can understand from the study by Smith et.al [16] that Standard methods of public health used during SARS have been successful and included aggressive case identification, case isolation, contact tracing and quarantine of all contacts, social distancing and group quarantine. However, the success of these interventions would largely depend on the differences between the SARS and COVID-19. Clear variations are emerging, such as COVID-19 has a higher transmissibility than SARS, and far more patients with COVID-19 rather than SARS have mild symptoms that adds to transmission, as these patients are still ignored and are not isolated. Due to the degree of population dissemination, conventional public health interventions do not stop all human-to-human transmission. Therefore, the focus should be more on mitigation rather than containment measures. Lack of vaccination and effective care has left us with only option of isolation and quarantine, social distancing, and group containment strategies. These are the only viable public health resources to manage transmittable diseases from one person to another. Though there is a huge socio-economic impact of social distancing, still it is much lesser than the long-term cost due to non-containment of the virus.

Additionally, ethical considerations, unbiased approach of the health authorities and political leaders, economic impact considerations, attention towards the vulnerable population and strict actions against individuals who deny treatment; are few key measures suggested by Lewnard et.al [Error! Reference source not found.] that will contribute to the success or failure of the social distancing strategy against the spread of COVID-19 in any country.

However, this is the theoretical or in vitro concept that is being studied in various research papers. But it is upon the compliance of the general population that decides the success of such mitigation strategies in any country. For a better understanding, this study is conducted to get a speculation on the perceived effectiveness of social distancing strategies to combat against the COVID-19 infection in India.

Materials and methods

A primary research study was done using an online questionnaire created using google forms. An observational study is done. Purposive sampling (Non-probability sampling) was done to select the respondents (Calculation for sample size selection is shown in Equation.1). The respondents were forwarded a survey of 14 questions including a consent form to participate in the survey, using mail ids and social media platforms. The survey link was opened on 21 April, 2020 at 10.00 GMT +5.30 and closed on 26 April, 2020 at 12.00 GMT +5.30. As many as, 413 participants responded to the survey from different states of India, the mapping of which is shown in the Figure.5. The participants were above 18 years of age who can give consent for themselves with an ability to read and write English having an active internet connection and having a basic knowledge about healthcare.

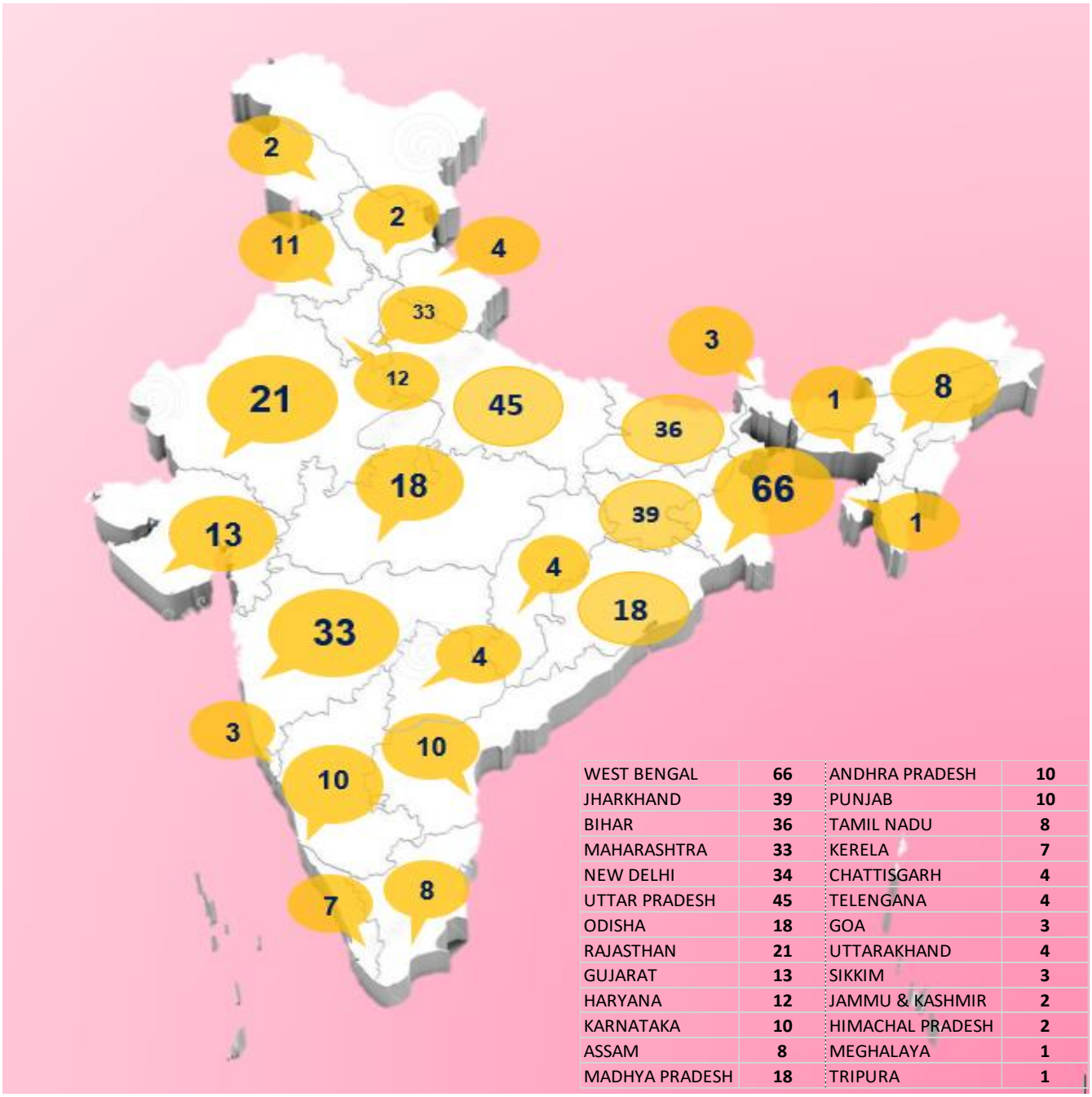
Equation 1: Shows sample size (N) calculation for a target population of more than 10,000, taking confidence interval (C.I.) as 95% i.e. Z value= 1.96; the proportion of population, p = 50% or 0.5, q = (1-p); degree of accuracy, d = 0.05. The number of non-respondents were taken to be 10% for the calculation of final sample size (N_f).

$$N = \frac{z^2 pq}{d^2} = \frac{(1.96)^2 \; X \; 0.5 \; X \; (1-0.5)}{(0.05)^2} = 384.14$$

$$N_f = \frac{\text{Effective sample size}}{(1-\text{non response rate anticipated})} = \frac{384.14}{1-0.1} = 427$$

To begin with, the questionnaire consisted of a consent form so that the participants willingly participate in the survey. After giving the required consent, the participants were presented with questions regarding their demographic details. In the next section, the questionnaire focused on three aspects i.e., awareness, attitude and barriers of social distancing measures to fight the spread of novel coronavirus in India. There were 4 multiple choice questions related to the awareness, followed by 2 questions on attitude that were rated in 3-point Likert scale, and 3 questions were put in barrier section, where one question was of multiple choice and remaining two were rated in 3-point Likert scale. After the data was collected, analysis and interpretation of the data were done using basic statistical tools like mean, percentage distribution, and standard deviation. Analysis was done with the help of Ms Excel spreadsheets. After obtaining the key findings, the data was represented using Graphical representations.

Websites of WHO, CDC and MoHFW were taken as reference for the secondary data used in the study regarding the number of cases of coronavirus infection, globally and nationwide [1] [8] [Error! Reference source not found.]. Review of literature for conducting the study were searched with the help of PubMed, ResearchGate and Google Scholar. The research papers dated from 2019 – 2020 were studied to understand the impact of social distancing strategies on the number of cases of COVID-19 nationwide and globally.



WEST BENGAL	66	ANDHRA PRADESH	10
JHARKHAND	39	PUNJAB	10
BIHAR	36	TAMIL NADU	8
MAHARASHTRA	33	KERELA	7
NEW DELHI	34	CHATTISGARH	4
UTTAR PRADESH	45	TELENGANA	4
ODISHA	18	GOA	3
RAJASTHAN	21	UTTARAKHAND	4
GUJARAT	13	SIKKIM	3
HARYANA	12	JAMMU & KASHMIR	2
KARNATAKA	10	HIMACHAL PRADESH	2
ASSAM	8	MEGHALAYA	1
MADHYA PRADESH	18	TRIPURA	1

Figure 5: Number of respondents of the survey from all over India – state wise distribution. Source. Original.

There were 413 respondents from 24 states and 2 union territories of India (Figure 5), in which 1 respondent did not agree to participate in the survey. The demographic details in the survey revealed that, out of the 412 participants, 52% were female and remaining 48% were male. The average age of the participants was 28.26 years and the median age of the same was 26 years. Most of the respondents were qualified with a graduate degree or more.

KEY FINDINGS

Awareness about the novel coronavirus and social distancing

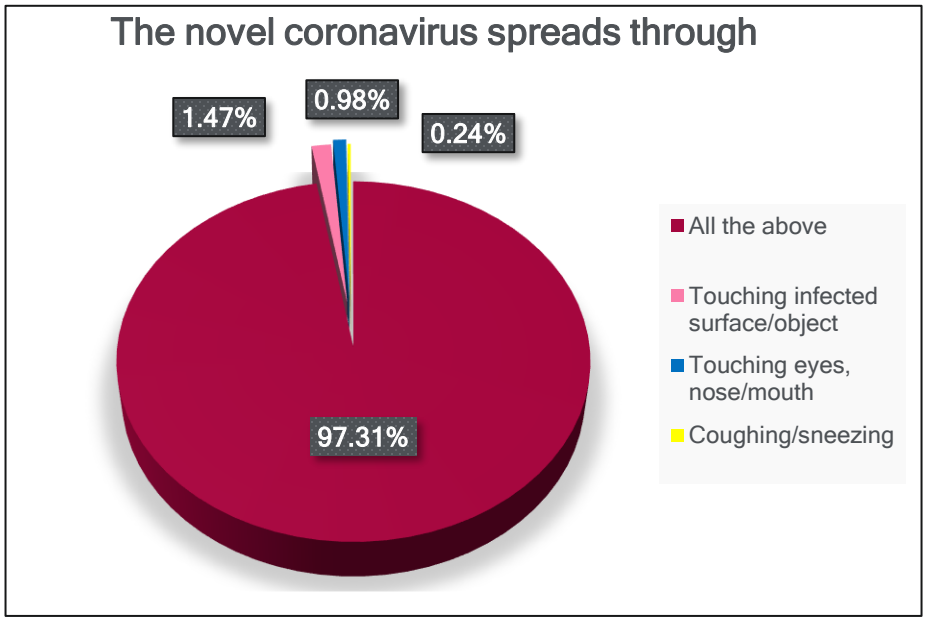


Figure 6: Awareness about the modes of transmission of novel coronavirus. Source. Original.

As per Figure 6, a majority (around 97 percent) of the respondents were aware of all the modes of transmission of the novel coronavirus infection i.e., coughing, sneezing, touching infected object or surfaces and touching eyes, nose or mouth. The remaining (around 3 percent) were aware of any one of the modes of spread.

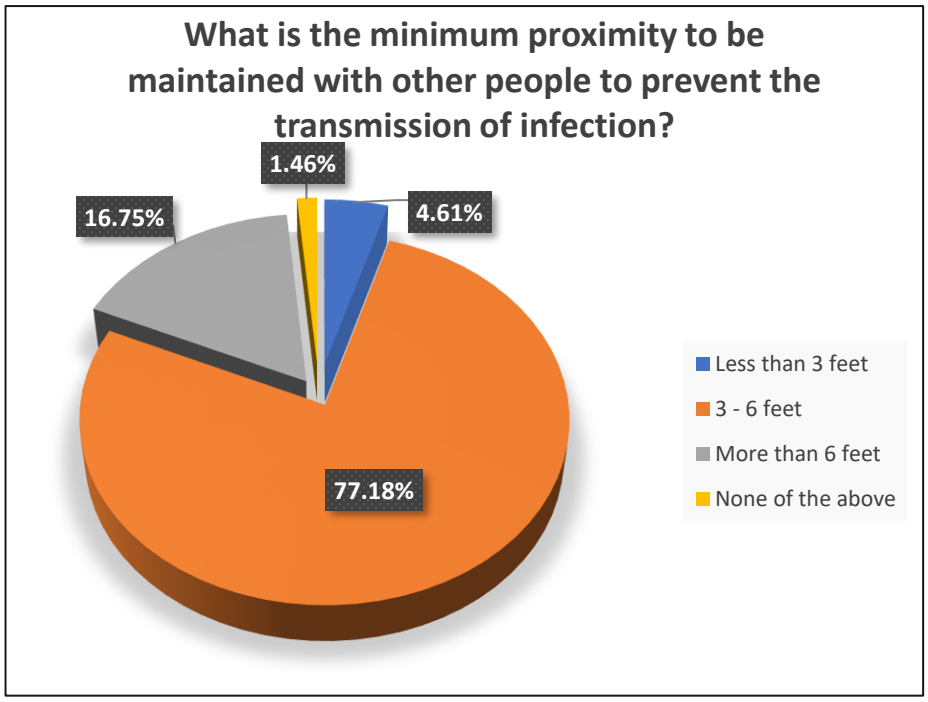
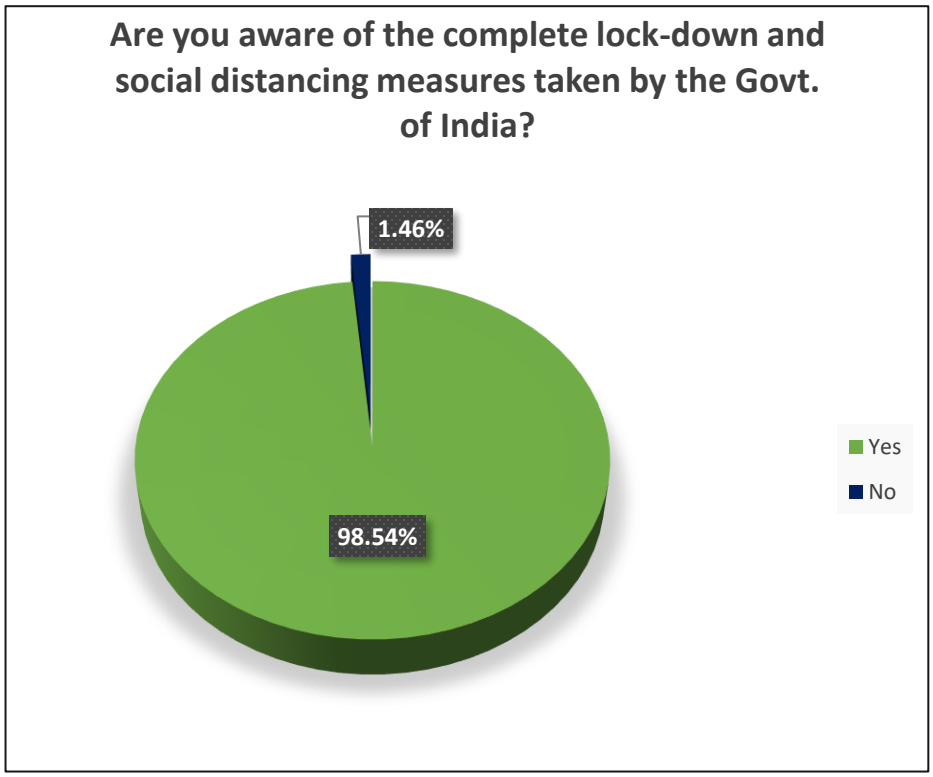


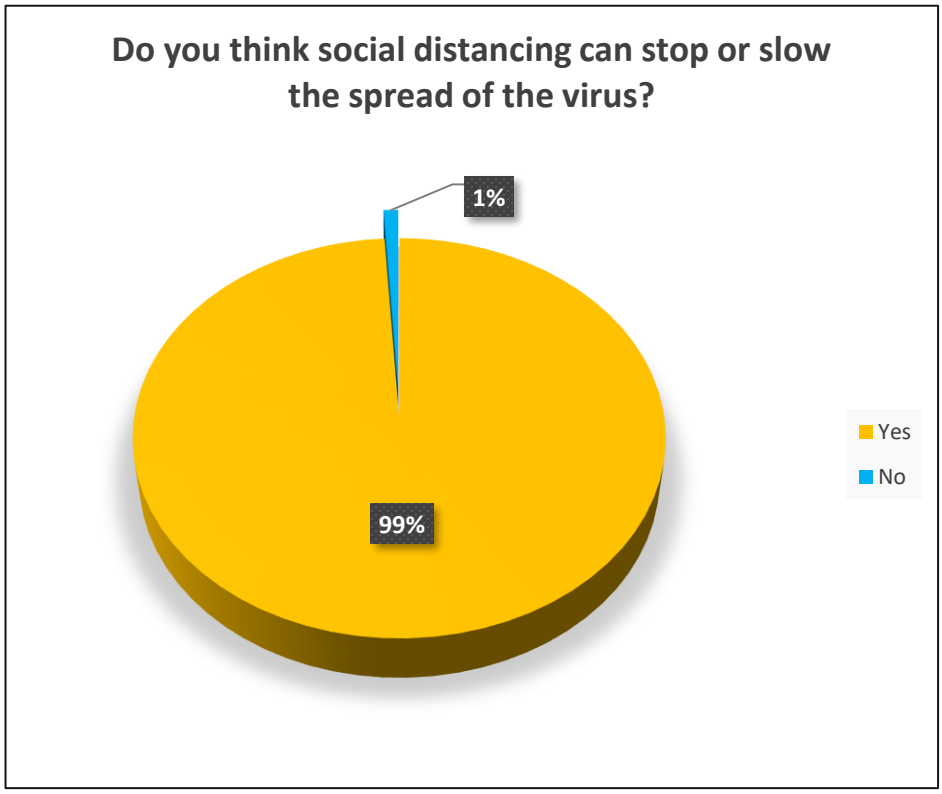
Figure 7: Awareness about the minimum proximal distance that should be maintained from one person to another to prevent the transmission of virus. Source. Original.

Figure 7 shows, around 77 percent of the respondents believed that the minimum proximity to be maintained from other people is 3-6 feet, whereas quite a few (around 17 percent) believed the same to be more than 6 feet. Very few of them (around 5%) answered it to be less than 3 feet. Rest of them were unaware of the minimum distance of proximity.



The majority (around 99 percent) of respondents were aware of the complete lockdown and social distancing measures being imposed by the Govt. of India since March 25, 2020 and rest 1 percent were not aware about these public health measures taken by the government as per Figure 8.

Figure 8: Awareness about the complete lockdown and social distancing measures taken by the Government of India. Source. Original.



In figure 9, the majority (around 99 percent) of the respondents believed that the social distancing measures are helpful in stopping or slowing the spread of the virus. Only a few (around 1 percent) respondents were in belief that social distancing has no effect in the act of spreading the virus.

Figure 9: Awareness about the effectiveness of social distancing in stopping or slowing the spread of the virus. Source. Original.

Attitude towards following public health mitigation strategies

Table 1: Findings on attitude of the people regarding following social distancing measures are represented in percentages. Source. Original.

During this period of Lockdown and social distancing, rate the following that is most suitable to you as, how likely would you;			
	Most likely (in %)	**Sometimes (in %)**	**Not likely (in %)**
Isolate yourself if you have symptoms like fever or cough	95.15	3.64	1.21
Avoid going out if it is not absolutely necessary	92.23	6.80	0.97
Use mask/cloth to cover your face while going out	94.18	5.58	0.24
Wash your hands frequently with soap/ Use sanitizer	92.72	7.04	0.24
Order food from outside	5.10	14.56	80.34
Avoid partying/ visiting relatives or friends	84.46	3.40	12.14

According to the responses in Table 1, a majority (around 95 percent) of the respondents were aware to isolate themselves if they had any symptoms like fever or cough, whereas quite a few (around 4 percent) of them would prefer to practice this sometime, and remaining 1 percent denied to adopt such practices. In the next section, most of the respondents (around 92 percent) would follow the community lockdown and avoid going out except for necessary chores, while a few of them (around 7 percent) would sometime go out without any necessity, and as few as 1 percent would not follow such practices. Further in the next section, a majority (around 94 percent) respondents would use mask or cloth to cover face while going out, whereas a few (around 6 percent) of them would follow this sometime. Moving further, a majority (around 93 percent) of respondents would follow hand hygiene practices with as few as around 7 percent of respondents who would follow this practice sometimes. In the next section, most of the respondents (around 80 percent) would not order food from outside, while around 15 percent of them might order sometimes and 5 percent of them would very likely order food from outside. Lastly, a majority (around 85 percent) of respondents would follow social distancing and avoid partying or visiting relatives, while around 3 percent would prefer to do such sometime, and around 12 percent would never avoid doing such thing.

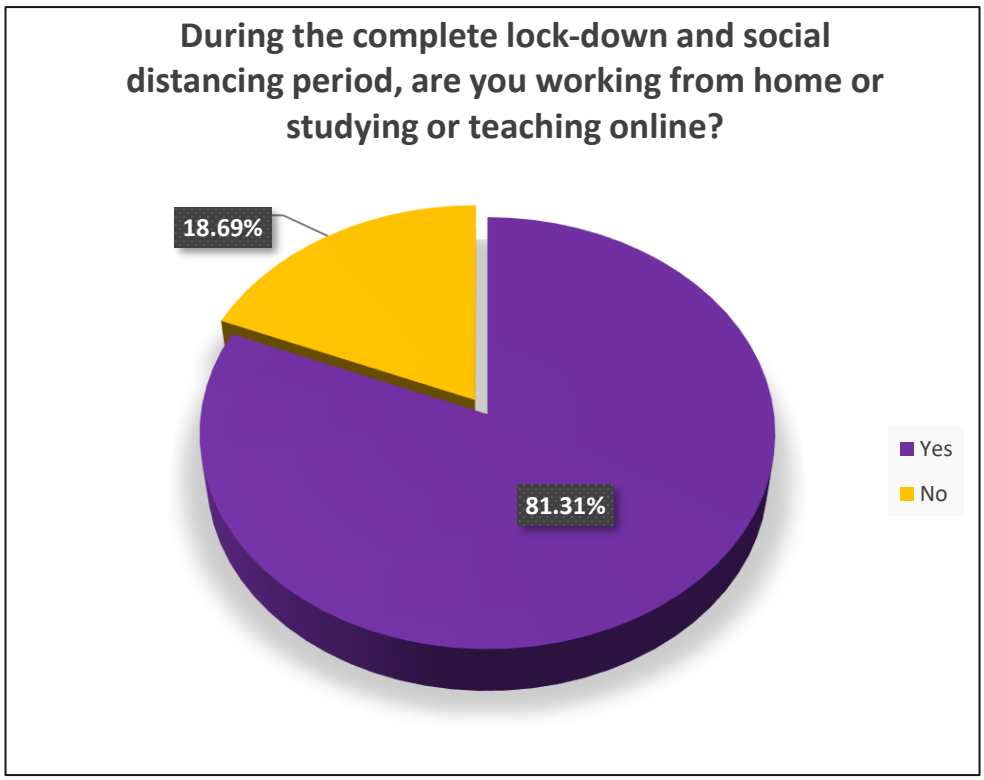
Perception towards effectiveness of public health measures

Table 2: Findings on the perception of the people about effectiveness of different public health mitigation measures including Social distancing and effective working of the frontline workers. Source. Original.

In your opinion, consider the following measures that has helped India to fight the spread of novel coronavirus and rate them with suitable answer.			
	Highly Effective (in %)	**Somewhat Effective (in %)**	**Not at all Effective (in %)**
Social Distancing	89.81	9.71	0.48
Case Isolation	92.48	6.31	1.21
Community lockdown	85.68	12.38	1.94
Work from home	80.83	17.96	1.21
Wearing mask	85.44	13.59	0.97
Covering face while coughing or sneezing	88.83	9.95	1.21
Washing hands	93.93	5.10	0.97
Effective working of the frontline workers like doctors, nurses, police, cleaning staff, and other essential service providers	94.91	4.85	0.24

As seen in Table 2, a majority of respondents have a perception that Social Distancing (around 90 percent), Case Isolation (around 93 percent), Community Lockdown (around 86 percent), Working from home (around 81 percent), Wearing masks while going out (around 85 percent), Covering face while coughing or sneezing (around 89 percent), washing of hands frequently (around 94 percent), and effective working of the frontline workers (around 95 percent) are effective measures that have helped India to fight the spread of novel coronavirus so far. Quite a few of the respondents, i.e., around 10 percent, 6 percent, 12 percent, 18 percent, 14 percent, 10 percent. 5 percent and 5 percent, perceive that social distancing, case isolation. Community lockdown, work from home, wearing mask, covering face while coughing or sneezing, washing hands, and effective working of frontline workers respectively, are somewhat effective strategies. Rest are in opinion that these measures are not at all effective in fighting the spread of novel coronavirus in India.

Practicing social distancing measures



During this period of lockdown and social distancing, a majority (around 81 percent) of the respondents are working from home or studying or teaching online, as seen in Figure 10. Rest of them (around 19 percent) are not working, studying or teaching remotely.

Figure 10: Findings on the number of respondents who are in a remote work or study set up. Source. Original.

Barriers in practicing social distancing measures

Table 3: Findings on barriers in practicing social distancing measures. Source. Original.

Rate the following with the most suitable answer.			
	Mostly (in %)	**Sometimes (in %)**	**Never (in %)**
Are you facing challenge or difficulty with the all online set up?	37.14	49.51	13.35
Do you think you are facing loss in career or academics?	45.39	42.72	11.89
Do you buy essential stuffs all at once to stock at home?	16.51	49.51	33.98
Do you face shortage of necessary products and services?	52.43	32.77	14.80

As per Table 3, the challenge from all online set up is faced more often by around 37 percent of the respondents and sometimes by around 50% of the respondents, whereas rest (around 13 percent) of them do not face any such problem. Moving to the next section, quite a few (around 45 percent) of the respondents are facing loss in their career or academics during the lockdown, whereas quite a few more (around 43 percent) face the same sometimes, and the rest (around 12 percent) are not facing any such loss. Furthermore, few (around 17 percent) of the respondents most likely do panic buying, and quite a few (around 50 percent) do such occasionally, while rest (around 34 percent) of them never indulge in panic buying of essential stuffs. More than 50% of respondents accepted to been facing shortage of necessary products and services like medicines, masks, sanitizers etc., while few (around 33 percent) of them face such problem sometimes, and rest (around 15 percent) of them do not face any such shortage.

Mental Health status in midst of social distancing

Table 4: Findings on mental health of people during the community lockdown and social distancing represented in percentages. Source. Original.

How often do you feel the following way since the lockdown and social distancing has started? Rate the following with the most suitable answer.			
	Always (in %)	**Sometimes (in %)**	**Never (in %)**
Anxious or panic about current situation in India	43.20	43.69	13.11
Irritated or agitated mostly during the day	37.14	39.32	23.54
Short-tempered on most occasions	35.68	32.77	31.55
Lonely or abandoned	37.38	32.28	30.34
Sleeplessness worrying about future	46.84	30.83	22.33

As per the findings of Table 4, quite a few (around 43 percent) of the respondents always feel anxious or panicked about the current situation in India, and quite a few more (around 44 percent) feel such sometimes, while the rest (around 13 percent) never feel the similar way. Moving further, quite a few i.e., around 37 percent and 39 percent of the respondents respectively, are always or sometimes irritated or agitated during most part of the day, while the rest (around 24 percent) of them do not feel the similar way. Around 36 percent of the respondents always become short tempered on most occasions, and 32 percent of them feel the same sometimes, while the rest (around 30 percent) have no such feeling. Lastly, more than 46 percent respondents are always sleepless worrying about future, with a few (around 31 percent) of them facing the same situation sometimes, while remaining (around 22 percent) respondents do not have any such problem.

DISCUSSION

COVID-19, the all new pandemic, has put the world in the face of uncertainty and unknown challenges. Pandemics are dual killers; they kill people from the inside out. The deaths, economic losses, fallen share markets etc are visible impacts. But what can't be seen is the huge mental burden it puts on the lives of people. Lack of strategic preparedness plans, lack of awareness about the infection, and nonchalant attitude, keeps on adding to the burden of the disease. Countries across the world, have adopted mitigation measures like social distancing in the battle against COVID-19. But it is crucial to analyse its perceived effectiveness by the ones who have to follow this. This study focusses on studying the awareness, attitudes and barriers of the social distancing to understand the perceived effectiveness of such measures in India.

Studying the responses in the awareness section, there is an evidence that a majority of respondents were aware of the current situation of COVID-19 infection and the mitigation adopted in India like community lockdown and social distancing measures. They were also aware of the modes of transmission and the minimum proximity of distance that should be maintained to prevent the transmission of infection, i.e., up to 3-6 feet or more than that. A majority of respondents believed that social distancing is helpful in stopping or slowing the spread of the virus.
Being aware of the infection, its transmission and the importance of the mitigation strategies, are crucial in shaping up the behaviour of the people in this fight against COVID-19. It will also instigate readiness to follow instructions given by the health authorities as preventive strategies.

In the next section, the aim was to assess whether the awareness in the people are also becoming a part of their attitude. More than 90% of the respondents would go into self-quarantine if at all they had any symptoms of fever and cough. A majority of them would avoid going out of their homes if at all the reason was not that necessary. Additionally, if they had to go out for some reason, more than 90% would prefer to use mask or cloth to cover their face. Hand hygiene practices were also seen to be followed by many. In the other hand, more than 80% of the respondents would avoid ordering food from outside and would rather prefer to eat home cooked meals, with a majority of them preferring to avoid visiting any relative or friend.

Secondly, more than 80-90% respondents agreed that social distancing measures like case isolation, community lockdown, work from home, proper use of masks, protection while coughing or sneezing, frequent hand hygiene practices with effective working of frontline workers like doctors, nurses, ASHAs, policemen, cleaning staff etc.; play a great role to combat the transmission of infection of COVID-19. Since the dawn of the complete lockdown, our prime minister and other health authorities are stressing to impart these healthy behaviours in our lifestyle, so

that we can protect ourselves from the risk of the infection. In every interaction of the prime minister with the countrymen, he has shown his gratitude towards the frontline workers, who are now termed as corona warriors by him.

Despite this, violence against these corona warriors are also coming to light from some part of the country or the other. It's a matter of shame and also a great concern in a situation where the ratio of patients to the healthcare providers are already below the minimum set proportion and these frontline workers are risking their lives to break the chain of infection.

In midst of the increasing number of cases, and closure of school, colleges and other work organizations, social distancing and prolonged lockdown are brewing uncertainty, panic hoarding and a debridement of the mental health of the people.

So, in the last section, the study aimed to analyse the barriers that prevents the awareness to turn into attitude. As per the data, it was seen that a majority of people were facing challenge from the all online set up while working or studying online, sometime or the other.

Nearly 60% of the respondents are facing some kind of loss in their academics or career, mostly or sometimes. As many as 50% respondents accepted to have indulged in panic hoarding of essential stuffs sometime or the other. Additionally, for the same reason, more than 50% are facing shortage of the necessary products and services.

Taking a peek into the mental health scenario, more than 80% of the respondents were facing anxiety issues or getting panicked, some time or the other, seeing the current situation of India.

In addition, a majority of the respondents accepted to be irritated or short tempered or being sleepless worrying about the uncertainty of their future.

S.W.O.T. Analysis of Social Distancing with respect to COVID-19

S — STRENGTHS

- Reduced incidence rates of cases infected with COVID-19.
- Reduced case fatality rates and adequate time to build herd immunity.
- Increased and focussed action from response groups and task forces.
- Rapid testing and proper surveillance in hotspot areas.
- Adequate time for contact tracing from confirmed cases.
- A barrier to community transmission by timely case isolation.
- Adequate number of frontline workers to give qualitative service to already infected cases.
- Healthcare and hospital facilities available to critical patients, which would have been a challenge to provide if there would have been a burst of new cases.

W — WEAKNESSES

- Negative impact on the mental health of the people.
- Panic buying or hoarding of essential goods and products resulting in scarcity and unavailability of these to the masses.
- Negative impact on country's economy and share markets if continued for a prolonged period.
- Large uncertainty regarding the relative risk of the infection and the case fatality ratios.
- Mild, asymptomatic and those who are in incubation period might go undetected, posing a risk of community transmission.
- Supply chains of essential products like masks, PPE kits, sanitizers, rapid testing kits etc. and the response group capacity might get exhausted in long run.

O — OPPORTUNITIES

- Social distancing and lockdown measures might contain the spread of the infection, resulting in limited transmission of the infection.
- The limited transmission can revive the country's economy much faster and open doors for exports in trade.
- Adequate time for increased surveillance and extensive research work, provides opportunity to develop vaccines and medicines at a faster rate with a better knowledge about the novel coronavirus.

T — THREATS

- Social, economic and political disruption of country if social distancing and lockdown measures ar continued for a prolonged time.
- Huge setback to the globalised trade sector involving several countries most of which are badly affected by the pandemic.
- Imports from countries like China who have not yet proved to have contained the infection can pose a threat of resurgence of new cases.
- Increased fear leading to increased violence and stigmatisation against frontline workers.
- A deteriorated economy will increase poverty, unemployment and hunger, leading to increased crimes and violence.

Conclusion

Social distancing, isolation and community lockdown, as we see them, have proven to be of great impact to prevent the rate of transmission and the risk of new infection. But we can't deny their mental health impacts as well, as they bring with them fear, panic, uncertainty, and challenges about the present and the future scenario. People are getting inactive, facing the fear of getting infected, worrying about the economic loss, feeling lagged behind in their lives, getting uncertain about their future plans and many more.

Additionally, there has been numerous instances of violence and stigmatization against the frontline health workers. The stigma is also associated with the infected patients or those who have been quarantined at some point of time. This scenario is a result of lack of awareness and acceptability of strategies adopted by the govt. with an added fear of the unknown. Though the Govt. of India and health ministries have adopted policies to protect the sovereignty and dignity of the health professionals **[Error! Reference source not found.]**, the problem lies in the perception of the general population. The attacks on health workers who are engaged in collection of samples are mere examples of fear driven actions. This calls for strategies that are people friendly with proper counselling before proceeding towards implementing a strategy. A stepping stone towards such strategy is the launch on the Arogya Setu app by the Govt. of India. Though, its reach is limited to people owning smartphones and not to the poor, underprivileged masses.

Social distancing, not only, prevents the transmission of infection, it also provides the health workers with the precious time to do rapid testing of more and more people and provide adequate care to the patients who are infected with the virus. Hence, the Govt. should not rush into lifting the lockdown just by looking into the negative impacts. The positive impacts are long run and have already holded the explosion of infection in India so far.

The key issues identified are; stigmatization and violence, lack of awareness, panic buying and questionable mental health status. So, it is proposed that the Govt. should take the following steps in addition to the public health mitigation strategies;

- Increasing awareness of the general public about the COVID-19 infection through IEC materials, short films, radio spots etc., about how it spreads and most importantly, how it does not, so that stigmatization can be prevented.
- Proper counselling of the people prior to sample collection and testing visits. Help of media, renowned stars, dharma gurus or local people can be taken to increase the acceptability in the masses.
- Proper and adequate availability of necessary products and services to avoid panic buying and unnecessary hovering.
- Effective intervention strategies and encouraging people to remain fit and active to fight against their deteriorating mental health.

ABBREVIATIONS

WHO	World Health Organization
SARS CoV	Severe Acute Respiratory Syndrome – coronavirus
CEST	Central European Summer Time
GMT	Greenwich Meridian Time
S.W.O.T.	Strengths, Weaknesses, Opportunities, Threats
MoHFW	Ministry of Health and Family Welfare
ICMR	Indian Institute of Medical Research
C.I.	Confidence Interval
CDC	Centres for Disease Control and prevention
ASHA	Accredited Social Health Activist
GOVT.	Government
IEC	Information Knowledge Communication

REFERENCES

1. Coronavirus Disease (COVID-19) - events as they happen. (n.d.). Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen
2. (n.d.). Retrieved from https://covid19.who.int/
3. Advice for public. (n.d.). Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public
4. Singh, Rajesh & Adhikari, Ronojoy. (2020). Age-structured impact of social distancing on the COVID-19 epidemic in India (Updates at https://github.com/rajeshrinet/pyross). Retrieved from: https://www.researchgate.net/publication/340209224
5. Nunes-Vaz, R., 2020. Visualising the doubling time of COVID-19 allows comparison of the success of containment measures. Global Biosecurity, 1(3), p.None. doi: 10.31646/gbio.61
6. Ramani, S., & Singaravelu, N. (2020, April 18). Data: How fast are India's coronavirus cases doubling since the lockdown? Retrieved from https://www.thehindu.com/data/data-how-fast-are-indias-coronavirus-cases-doubling-since-the-lockdown/article31371220.ece
7. (n.d.). Retrieved from https://en.wikipedia.org/wiki/Template:2019percentE2percent80percent9320_coronavirus_pandemic_data
8. (n.d.). Retrieved from https://www.mohfw.gov.in/
9. (n.d.). Retrieved from https://www.icmr.gov.in/
10. #StaySafe, P. I. B. I. #S. H. (2020, April 23). Comparative analysis of confirmation rate of many #COVID19 affected countries from the day of confirmation of 400th case: #India seems to be doing well with the focused strategy: Media briefing on Status of #Covid_19india pic.twitter.com/RIHYEJURnc. Retrieved from https://twitter.com/PIB_India/status/1253306424429408256?utm_campaign=fullarticle&utm_medium=referral&utm_source=inshorts
11. Wilder-Smith, A., & Freedman, D. O. (2020). Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. Journal of Travel Medicine, 27(2). doi: 10.1093/jtm/taaa020
12. Koo, J. R., Cook, A. R., Park, M., Sun, Y., Sun, H., Lim, J. T., Dickens, B. L. (2020). Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. The Lancet Infectious Diseases. doi: 10.1016/s1473-3099(20)30162-6

13. Milne, G. J., & Xie, S. (2020). The Effectiveness of Social Distancing in Mitigating COVID-19 Spread: a modelling analysis. doi: 10.1101/2020.03.20.20040055

14. Kissler, S. M., Tedijanto, C., Lipsitch, M., & Grad, Y. (2020). Social distancing strategies for curbing the COVID-19 epidemic. doi: 10.1101/2020.03.22.20041079

15. Ault, A. (2020, March 27). COVID-19 and the Science of Social Distancing. Retrieved from https://www.medscape.com/viewarticle/927613

16. Zhang, Y., Jiang, B., Yuan, J., & Tao, Y. (2020, January 1). The impact of social distancing and epicenter lockdown on the COVID-19 epidemic in mainland China: A data-driven SEIQR model study. Retrieved from https://www.medrxiv.org/content/10.1101/2020.03.04.20031187v1

17. Wilder-Smith, A., Chiew, C. J., & Lee, V. J. (2020, March 5). Can we contain the COVID-19 outbreak with the same measures as for SARS? Retrieved from https://www.sciencedirect.com/science/article/pii/S1473309920301298

18. Lewnard, J. A., & Lo, N. C. (2020). Scientific and ethical basis for social-distancing interventions against COVID-19. The Lancet Infectious Diseases. doi: 10.1016/s1473-3099(20)30190-0

19. (n.d.). Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/world-map.html

20. Prez gives assent to Ordinance making acts of violence against healthcare workers cognizable, non-bailable offence. (n.d.). Retrieved from http://newsonair.com/Main-News-Details.aspx?id=386447

21.

REPORT – II

HYPOTHESIS TESTING IN PUBLIC HEALTH

JOHN HOPKINS UNIVERSITY
(COURSERA)

COURSE INSTRUCTOR: JOHN MCGREADY
PHD, MS(BIOSTATISTICS)
BLOOMBERG SCHOOL OF PUBLIC HEALTH

A REPORT ON ONLINE COURSE

Authored by:

DR. NABANITA MUKHERJEE

POST GRADUATE STUDENT
MBA HEALTHAND HOSPITAL MANAGEMENT
IIHMR UNIVERSITY, JAIPUR
REGD. NO. 0937

Guided by: Mr. Arindam Das

ASSOCIATE PROFESSOR
IIHMR UNIVERSITY, JAIPUR

IIHMR UNIVERSITY

ACKNOWLEDGEMENT

With immense pleasure, I present to you my summer training report part - 2 on an online course from coursera conducted by John Hopkins University, USA titled as "**Hypothesis Testing in Public Health**" as a part of the 1st year academic curriculum of MBA Health and Hospital management course.

I thank all who in one way or another contributed in the completion of this report.
I would like to extend my profound thanks to my mentor "ARINDAM DAS", under whose expert guidance and critical supervision, I could create and complete this report.
I am also deeply thankful to the instructor of my online course "John McGready", who gave me meaningful knowledge regarding the course and simplified every topic with relevant examples.

Lastly, I thank all those who have directly or indirectly contributed and helped in completion of my report.

DR. NABANITA MUKHERJEE
MBA-HM-24
2019-2021
REGD NO. 0937

TABLE OF CONTENTS

KEY LEARNINGS

About the course

DESCRIPTION AND OBJECTIVES

- Biostatistics is a basic aptitude for each public health scientist since it gives a lot of exact strategies for separating significant ends from information.
- In this course of Hypothesis testing, I have figured out how to assess variability in samples and apply statistical testing strategies.
- Along the way of this course, I have performed estimations and translation of true information from the published scientific literatures.
- Themes that are incorporated here are; sample statistics, the central limit theorem, confidence intervals, hypothesis testing, and p-values.
- **Duration**: The course is of total 19 hours duration.
- **Learning methodology**: Lectures, PPT, Video, Assignments, Graded quizzes.

Objectives

1. To find out sample statistics, sampling distribution and central limit theorem.
2. To know about confidence intervals, the need for ample size of data, and ways to get around the need for ample size of data.
3. To study statistical hypothesis test, confidence intervals and, p-values.
4. To know about comparing proportions and incidence rates between two populations.

COURSE CONTENTS

WEEK 1

- Sampling distribution definition.
- Examples of Sampling Distribution with Single mean
- Examples of Sampling Distribution for a Single Proportion, Incidence rate.
- Estimating Sampling Distribution characteristics from single samples of data.

WEEK 2

- Confidence Intervals for single Population parameters.
- Confidence Intervals for Population Proportions and Incidence Rates
- Computing and interpreting confidence intervals for binary comparisons - Odds Ratio
- Confidence interval for population quantities of small sample.

WEEK 3

- Hypothesis Testing: The General Concept, and for Comparing Means Between Two Populations
- Hypothesis Testing: The paired approach
- Hypothesis Testing: The unpaired approach
- Brief note on p-value and types of errors

WEEK 4

- Hypothesis Testing: The z-test approach
- Hypothesis Testing:Chi-Square and Fisher's Exact Test Approaches
- Hypothesis Tests for Comparing Incidence Rates of more than Two Populations - ANOVA, chi-square and log scale tests.

Key Learnings

1 Sampling Distribution and Standard errors

❑ Definitions

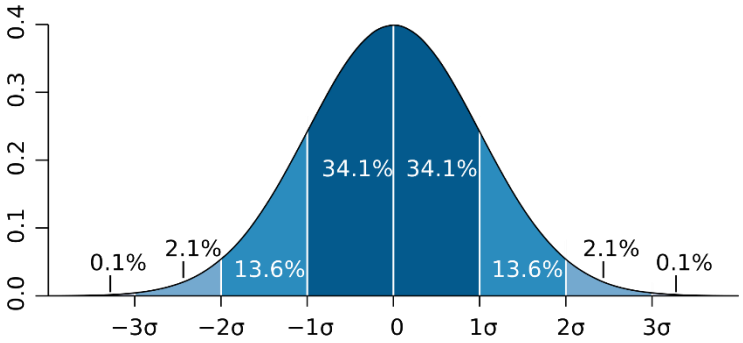
❖ Sampling distribution

The distribution of a sample data point "$\bar{X}$" derived from a random sample of sample size "n" from a sample population of "N" mean "μ" and variance "σ^2", is known as the sampling distribution of that sample data point.

$$\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

❖ Standard deviation

The standard deviation "σ " in a sample statistic is referred to as the measure of the amount of deviation from the predetermined set of values.



❖ Standard error

The standard deviation measured for a sampling distribution, is referred to as the standard error of that distribution, taking arithmetic mean for each sample as " $\bar{x}$ ".

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

❑ Examples of Sampling Distributions

➢ Sampling distribution for a single mean

- Here the histogram shows the 5,000-mean distribution, 5,000 mean sample heights, based on 5,000 samples of each size 20.
- Based on a random sample of N = 20, each point in this histogram is a single mean. And again, not the individual height values are the points in this histogram, but sample means each based on 20 individuals.
- The mean sample distribution tends to be **symmetrical and bell-shaped**. The distribution now takes the same form, based on 50 persons each, and based on 150 individuals each.
- Outcome: The mean across all sample sizes remains the same, whereas the variability or standard deviation tends to decrease as the sample size increases and the sample average means, irrespective of the sample size, that each mean is close to or equal to the underlying population mean.

Table 5. A sample population of 5000 with sample mean heights and standard deviation each of sample size 20, 50 and 150. Source and credit. Coursera

Sample size	Mean of 5000 sample means	S.D. of 5000 sample means
n = 20	167 cm	0.57 cm
n = 50	167 cm	0.35 cm
n = 150	167 cm	0.20 cm

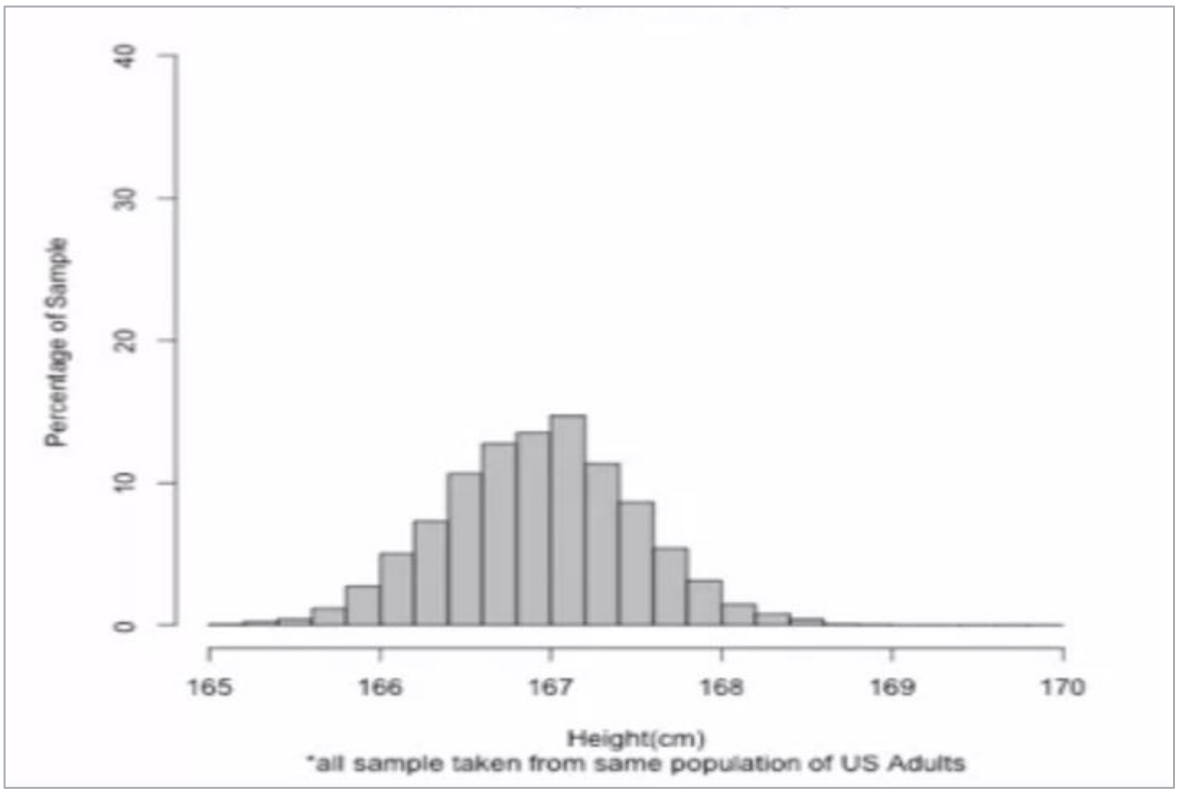


Figure 3. A histogram showing the 5000 sample mean distribution, 5,000 sample mean heights, based on 5,000 samples each of size 20. Source and credit. Coursera

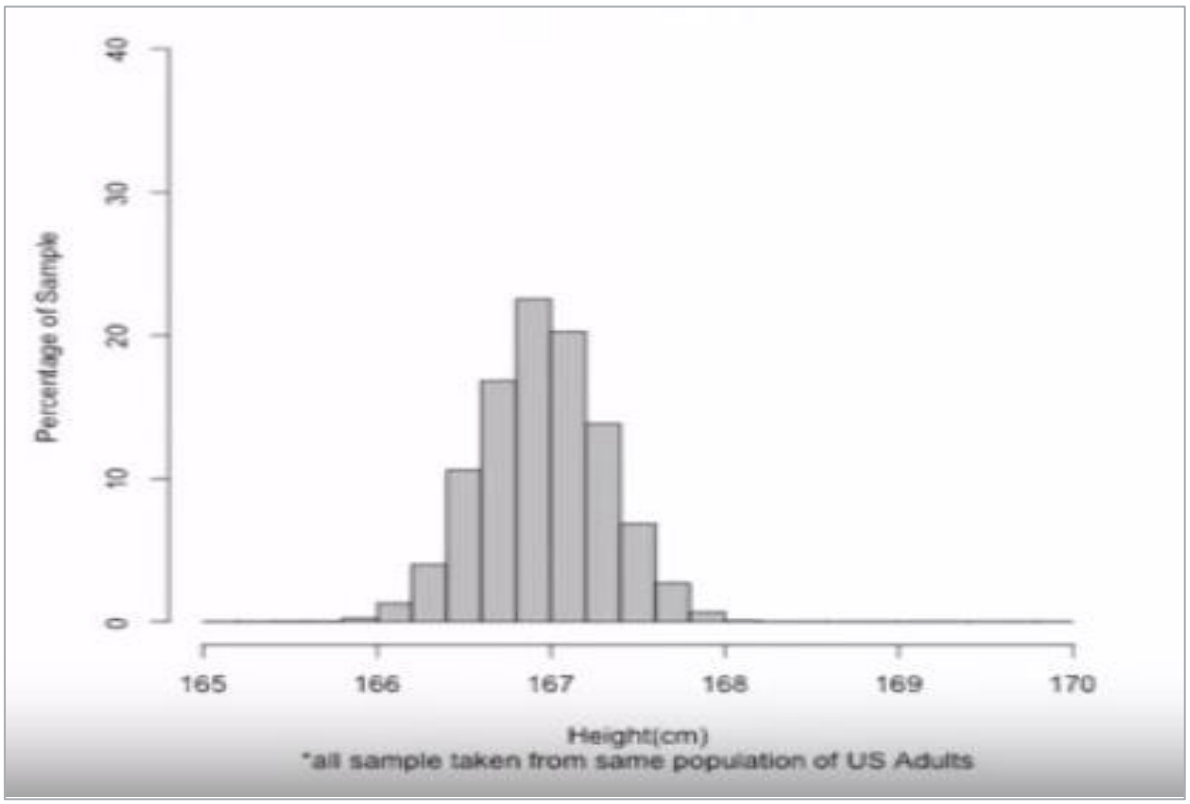


Figure 4. A histogram showing the 5000 sample mean distribution, 5,000 sample mean heights, based on 5,000 samples each of size 50. Source and credit. Coursera

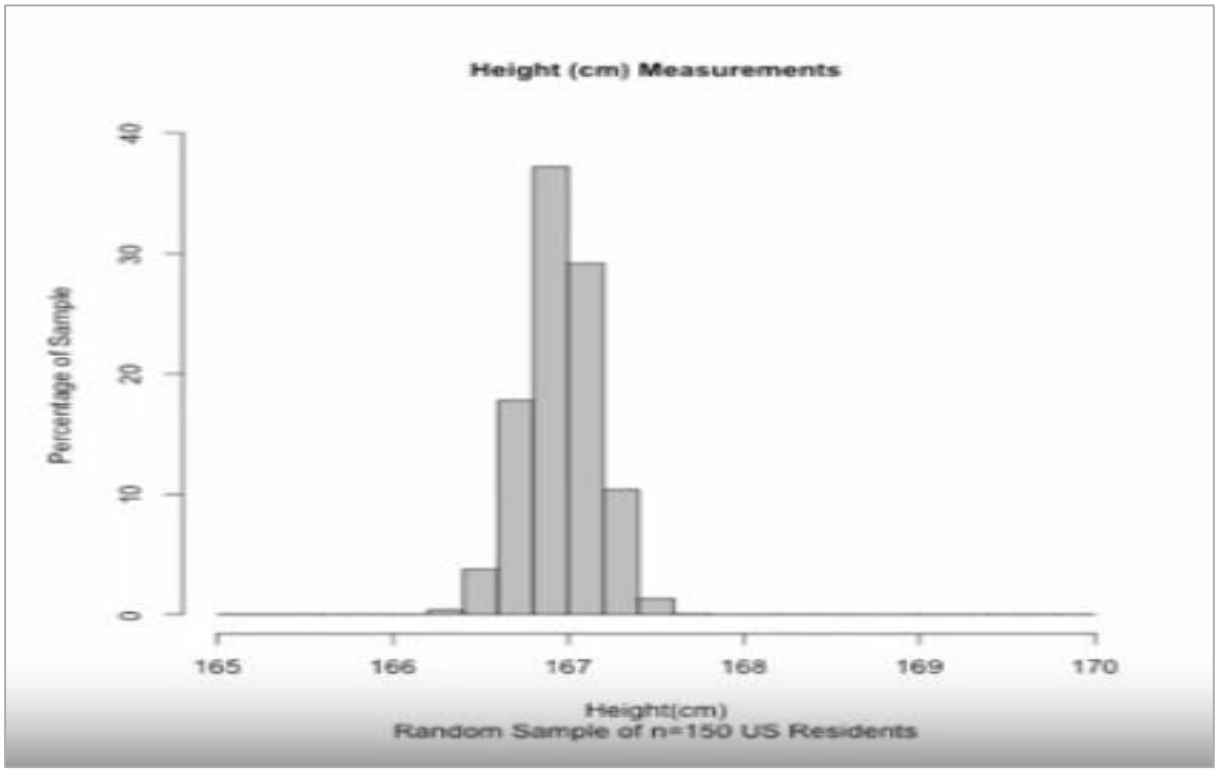


Figure 3. A histogram showing the 5000 sample mean distribution, 5,000 sample mean heights, based on 5,000 samples each of size 150. Source and credit. Coursera

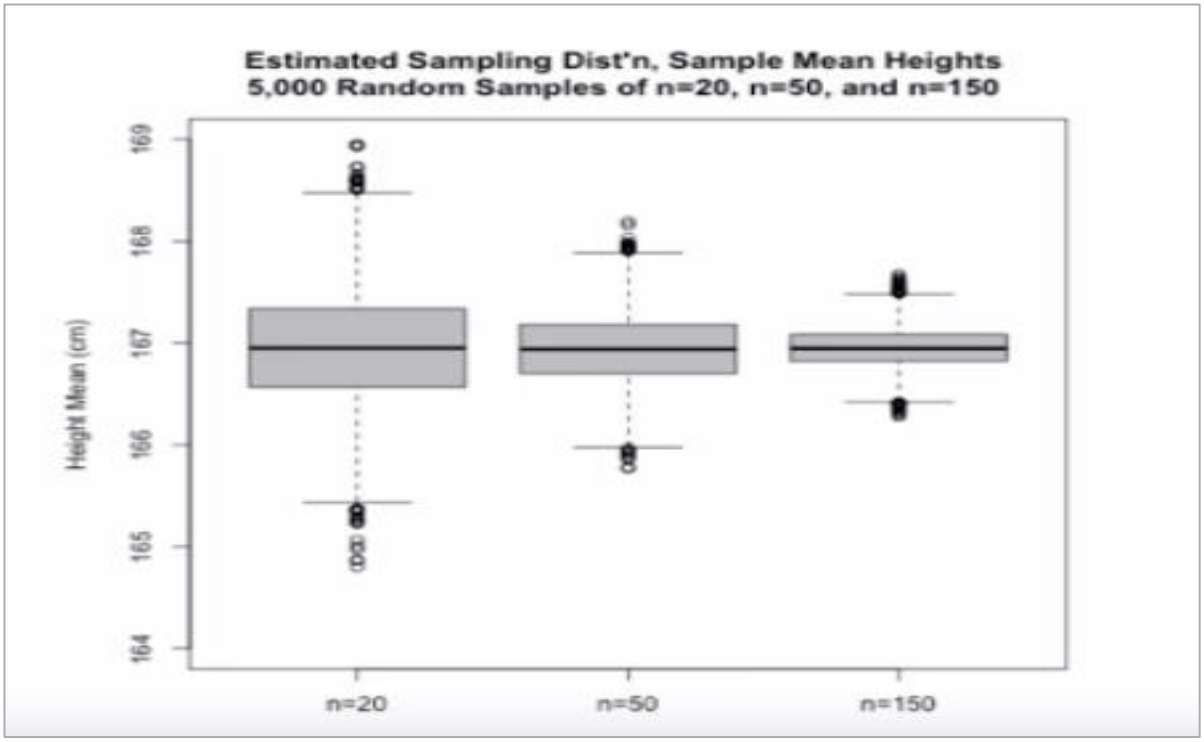


Figure 4. A box plot showing the 5000 sample mean distribution, 5,000 sample mean heights, based on 5,000 samples each of size 20, 50 and 150. Source and credit. Coursera

➢ Sampling Distribution for a Single Proportion

- Taking an example of poverty in Baltimore City in 5,000 students in a class.
- A random sample of 50 Baltimore City residents was taken, and the proportion in their sample who were living below the poverty line were computed.
- Same was repeated with a sample size of 150 and then 250. All 5,000 p-hats each based on 50 persons, 150 persons and 250 persons were plotted in histograms and box plots.
- Outcome: The average was similar as 25% for the means in all the three cases and the standard deviation for each ample proportion varied with increasing sample size. In other words, the larger the sample the proportion is based on, the less variability there is in proportions across the samples.

Table 6. An example of random sample resident of 50, 150 and 250 Baltimore City were taken, and the proportion in their sample means and standard deviations of residents living below the poverty line. Source and credit. Coursera

Sample size	Mean of 5000 sample proportions	S.D. of 5000 sample proportions
n = 50	0.25	0.061
n = 150	0.25	0.035
n = 250	0.25	0.019

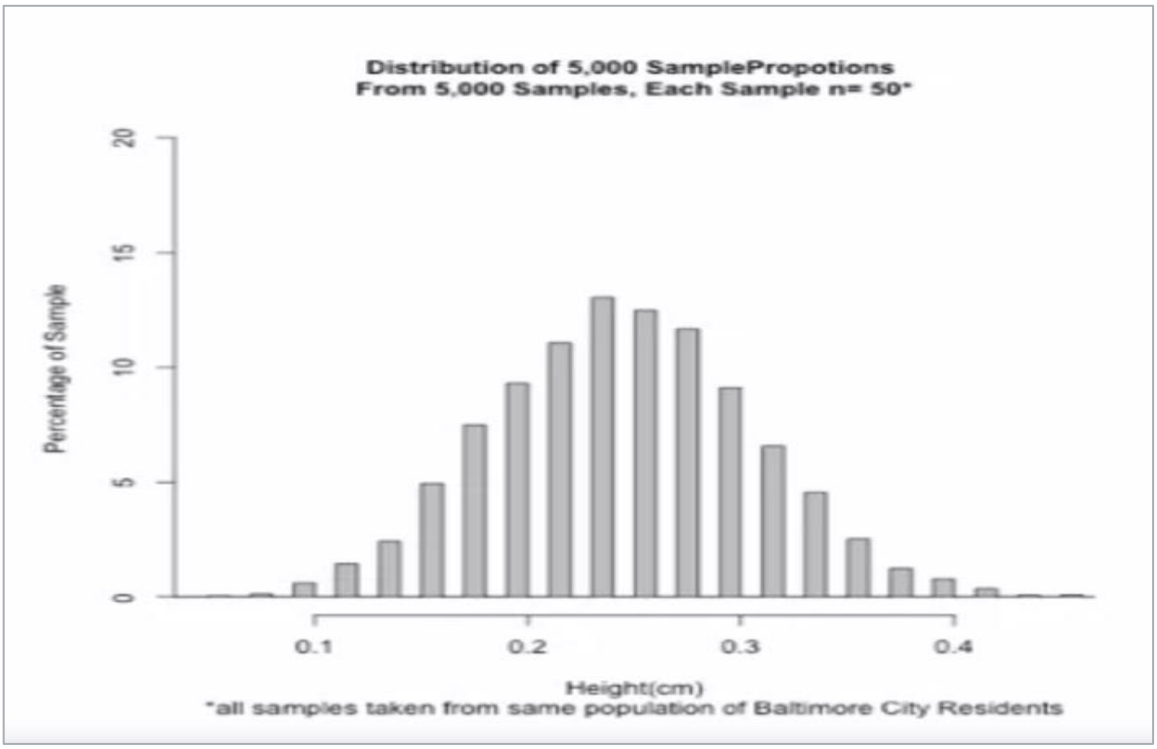


Figure 7 A histogram of distribution of 5000 sample proportions from 5000 samples, each of sample size 500. Source and credit. Coursera

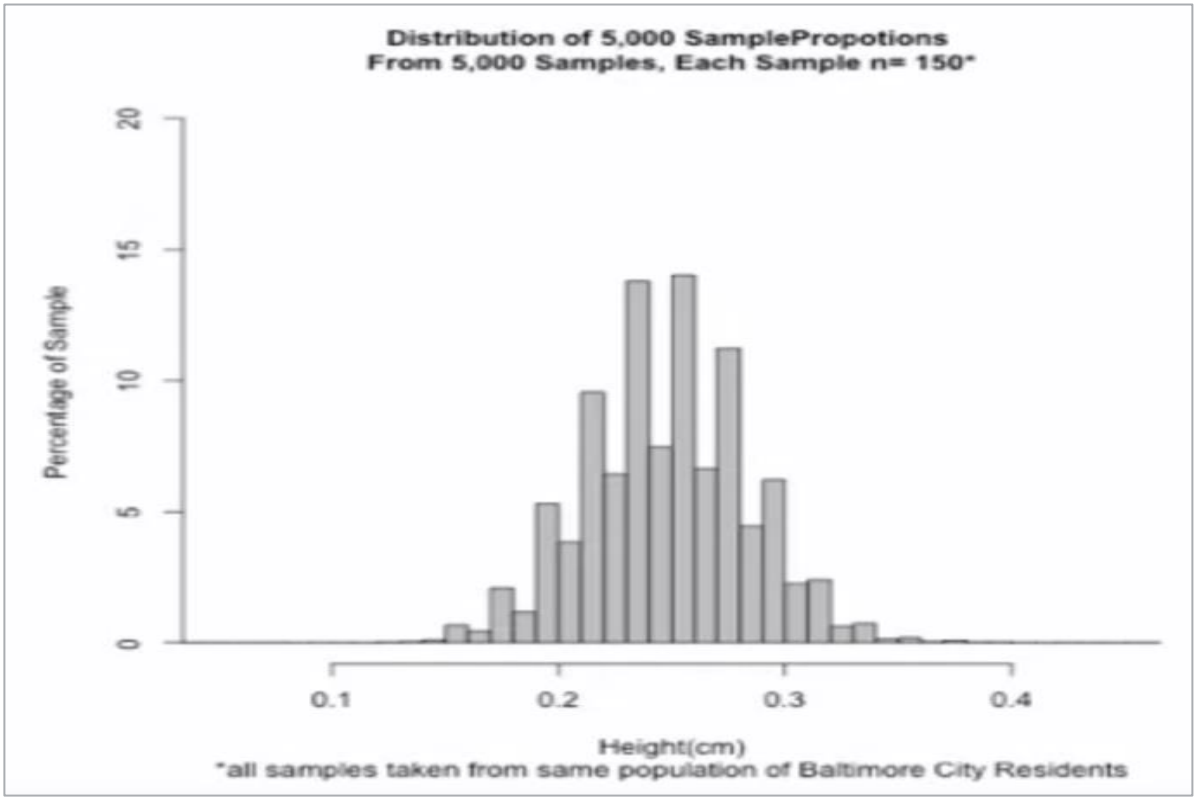


Figure 6 A histogram of distribution of 5000 sample proportions from 5000 samples, each of sample size 150. Source and credit. Coursera

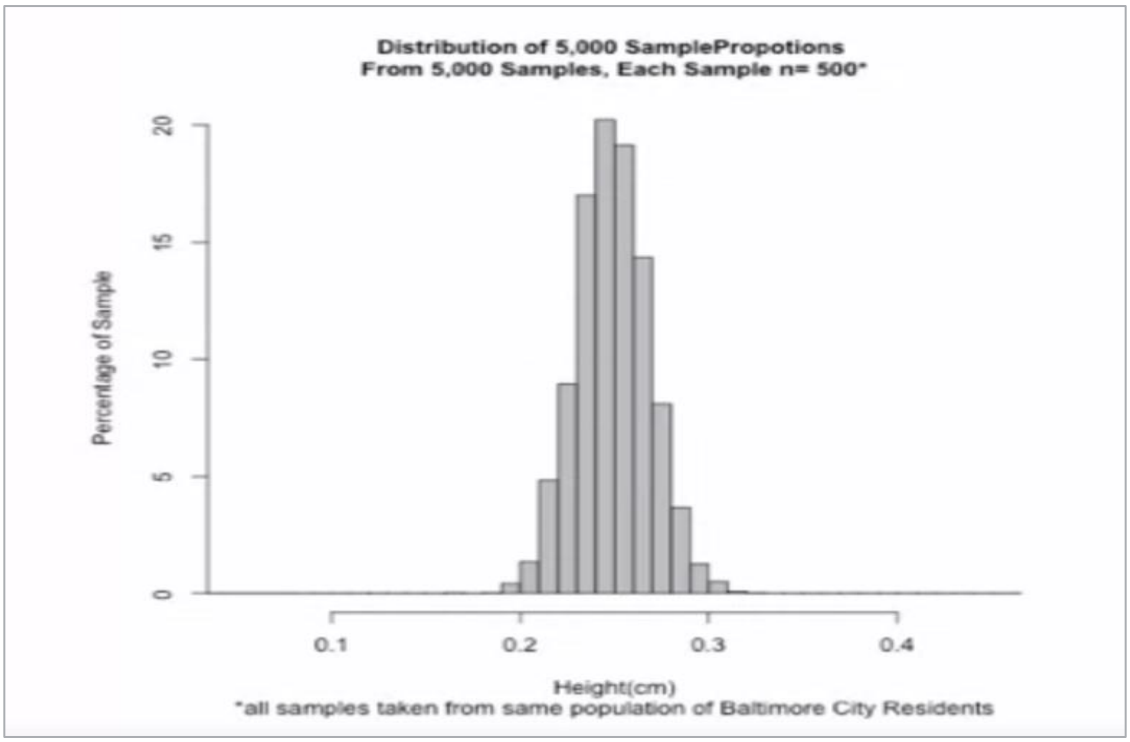


Figure 5. A histogram of distribution of 5000 sample proportions from 5000 samples, each of sample size 50. Source and credit. Coursera

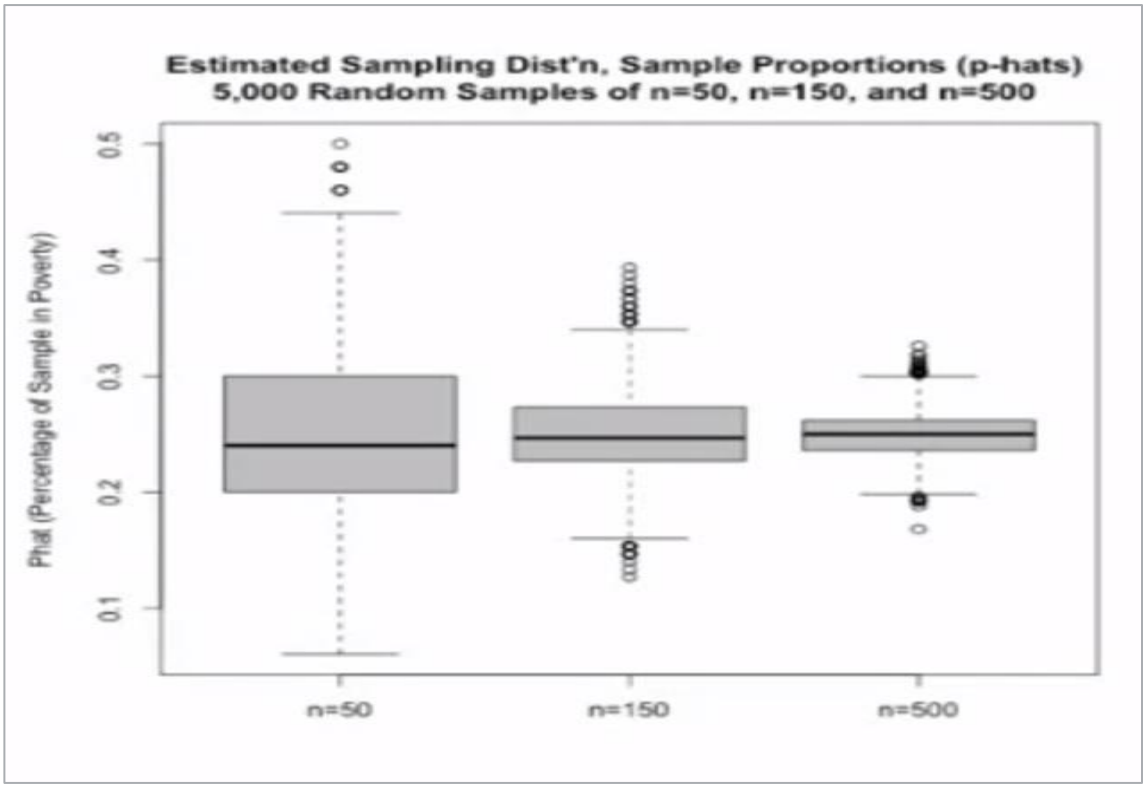


Figure 8 A box plot of distribution of 5000 sample proportions from 5000 samples, each of sample size 50, 150 and 500. Source and credit. Coursera

- Further, a data was collected on time to relapse for persons discharged from alcohol and drug rehabilitation facilities and data available on each participant clued their follow-up time, which is their time from discharged to relapse or censoring and whether or not the participant relapsed or was censored.
- Then to analyze the sampling behavior of a sample incidence rate using multiple sample sizes of size 50, a 100, and 500 respectively taken from the data set.
- Outcome: The median incidence rate of the estimator sample is similar if the rates are based on 50, 150, or 500 individuals, but the greater the sample on which each estimate is based, the variation in the incidence rates between samples tends to be lesser.

Table 7. A sample incidence rate using multiple sample sizes of size 50, a 100, and 500 respectively on time to relapse for persons discharged from alcohol and drug rehabilitation facilities and data available on each participant clued their follow-up time. Source and credit. Coursera

Sample size	Mean of 5000 incidence rates	S.D. of 5000 incidence rates
n = 50	128.6	22.7
n = 150	128.6	14.8
n = 500	128.6	8.1

❑ Estimating Sampling Distribution characteristics from a single sample of data

The central limit theorem states that;

if there is a population with mean μ and standard deviation σ and infinite number of random samples of different sizes are taken from the population, then the theoretical sampling distribution of the sample means will be approximately normally distributed.

On average, the average of our estimates will be the true population level value being estimated, and the variability in these estimates will be a function of both the variation in the individual values in the population, the standard deviation of the individual population values, and the size of the sample the statistics based on. This variability in sample statistics across multiple samples of the same size is called the standard error of the statistic.

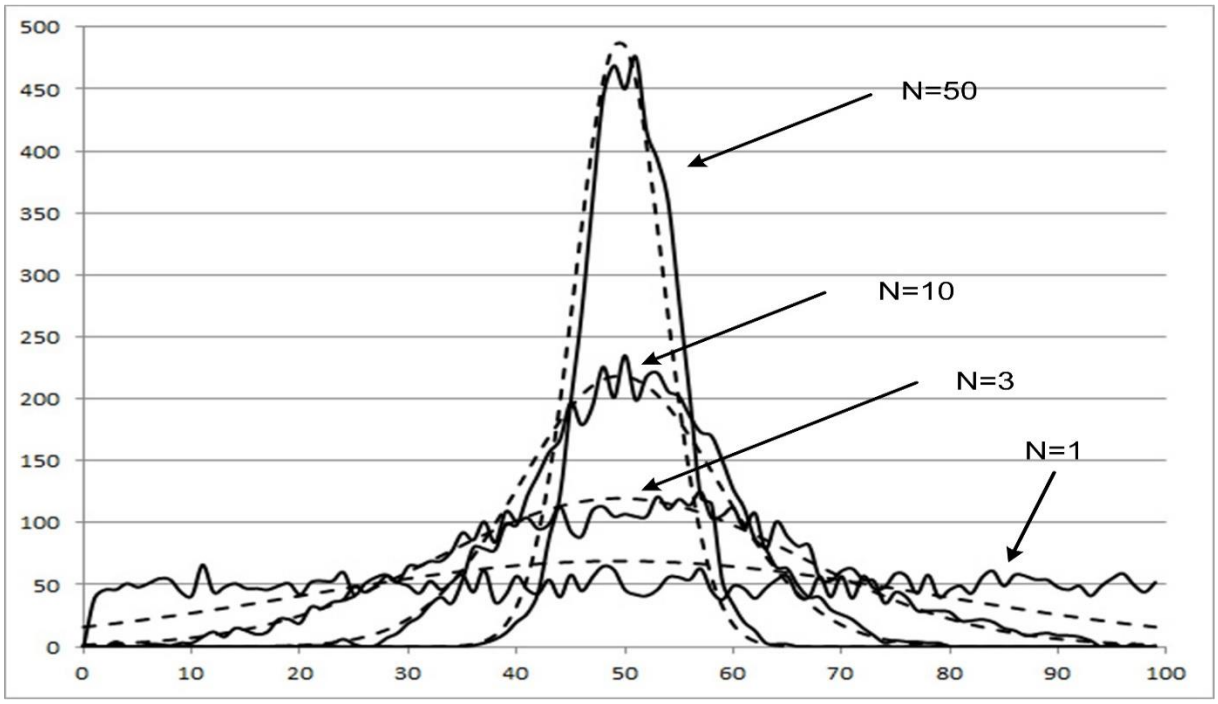


Figure 9. A sample graphical representation of central limit theorem with different sample sizes of n=1, 3, 10 and 50 respectively. Source and credit. UltraCAD Design, Inc

2 Confidence Intervals for Population parameters and

In statistics, a confidence interval refers to the probability that, for a certain proportion of times, a population parameter will fall between two fixed values. Confidence intervals in a sampling system are very useful to calculate the degree of uncertainty or certainty of that sample.

Most commonly used is the 95% confidence interval, and the other two being 99% and 90% confidence intervals.

The calculation of a 95% confidence interval for a population mean, indicated by "μ" based on a data from a randomly taken sample from a sample population, can be deducted by:

$$\bar{x} \pm 2\widehat{SE}(\bar{x})$$

where, $\widehat{SE}(\bar{x}) = \frac{\sigma}{\sqrt{n}}$

The level of confidence can be changed by adjusting the number of standard errors subtracted from and added to the sample mean.

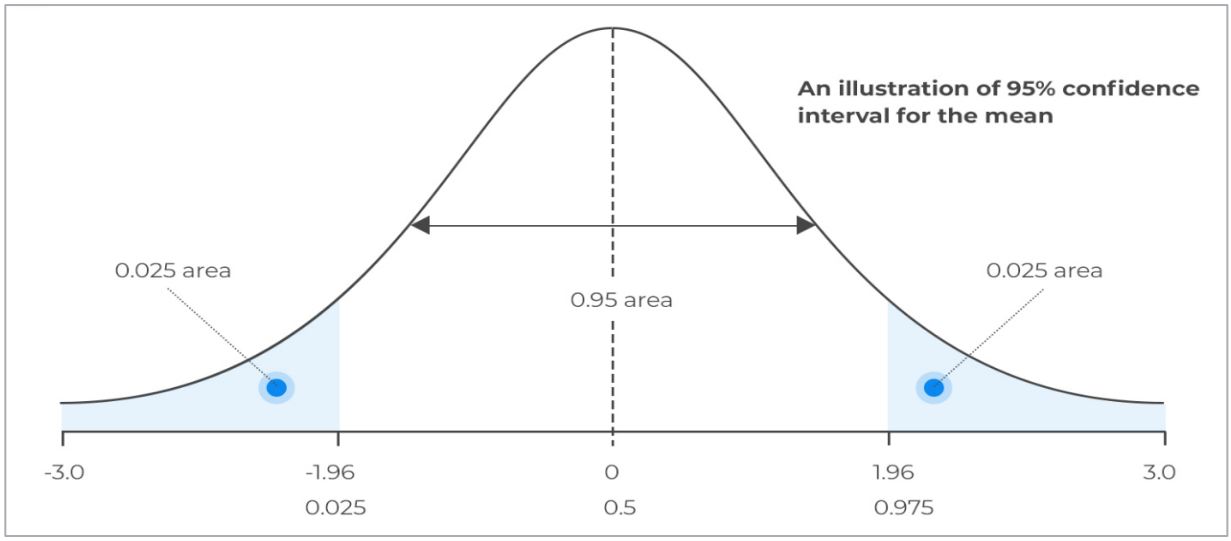


Figure 10. A graphical representation of 95 percent confidence interval of mean. Source and credit. AnalystPrep.

❑ Confidence Intervals – Proportions and Incidence Rates of Population

The equation for a 95% confidence interval for a population proportion "p" on the basis of a data from a randomly taken sample from a sample population, can be constructed by:

$$\hat{p} \pm 2\widehat{SE}(\hat{p})$$

where, $\widehat{SE}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

95% confidence intervals for a population incidence rate "IR" on the basis of a data from a randomly taken sample from a sample population followed up for a period of time, can be constructed by:

$$\widehat{IR} \pm 2\widehat{SE}(\widehat{IR})$$

where, $\widehat{IR} = \frac{\#events}{Total\ follow\ up\ time}$

$$\widehat{SE}(\widehat{IR}) = \frac{\sqrt{\#events}}{Total\ follow\ up\ time}$$

The standard error for this sample will quantify as the variance in sample proportions and incidence rates over random samples of the same size (from the same population), respectively.

❑ Confidence intervals - Binary comparisons of population (Odds Ratio)

The Confidence interval of a sample population for ratio-based measures of association with binary outcomes, both the relative risk and odds ratio, need to be computed on the natural log scale and then the results exponentiated (anti-logged) back to the ratio scale.

This can be further explained with the help of an example. Considering a random sample of 1,000 HIV positive patients from a citywide clinical population, and looking at the percentage of persons who responded to antiretroviral therapy based on their baseline CD4 count.

Table 8. An example of random sample of 1000 HIV positive patients showing percentage of persons who responded to antiretroviral therapy based on their baseline CD4 count. Source and credit. Coursera

Outcome	CD4<250	CD4≥250	Total
Responded	127	79	206
Not Responded	376	418	794
Total	503	497	1000

$\hat{p}_{CD4<250} = \frac{127}{503} = 0.25\ (25\%)$

$\hat{p}_{CD4\geq250} = \frac{79}{497} = 0.16\ (16\%)$

Relative Risk $\widehat{RR} = \frac{0.25}{0.16} \approx 1.56$

$\ln(\widehat{RR}) = 0.44$

Odds Ratio $\widehat{OR} = \frac{0.25/0.75}{0.16/0.84} \approx 1.75$

$\ln(\widehat{OR}) = 0.56$

I. $$\widehat{SE}\left(\ln\left(\widehat{RR}\right)\right) = \sqrt{\frac{1}{127} - \frac{1}{503} + \frac{1}{79} - \frac{1}{497}} \approx 0.13$$

$$\ln(\widehat{RR}) \pm 2\widehat{SE}\left(\ln\left(\widehat{RR}\right)\right) = 0.44 \pm 2(0.13) \rightarrow (0.18, 0.60)$$

This value needs to be exponentiated for Relative Risk,

$(e^{0.18}, e^{0.60}) \rightarrow (1.20, 1.82)$

RESULT- At the time of starting therapy, HIV-positive individuals with CD4 counts of less than 250 have a 56% higher risk or probability of responding to therapy compared to HIV-positive individuals with the count of CD4 cells of more than or equal to 250 at the beginning of therapy. These results estimate that this increase in response probability could be as small as 20 percent, and as large as 82 percent.

II. $\widehat{SE}\left(\ln\left(\widehat{OR}\right)\right) = \sqrt{\frac{1}{127} + \frac{1}{376} + \frac{1}{79} + \frac{1}{418}} \approx 0.16$

$\ln\left(\widehat{OR}\right) \pm 2\widehat{SE}\left(\ln\left(\widehat{OR}\right)\right) = 0.56 \pm 2(0.16) \rightarrow (0.24, 0.88)$

This value needs to be exponentiated for Odds Ratio,

$(e^{0.24}, e^{0.88}) \rightarrow (1.27, 2.41)$

RESULT - At the time of starting therapy, HIV-positive individuals with CD4 counts of less than 250 have a 56% higher odds of responding to therapy compared to HIV-positive individuals with the count of CD4 cells of more than or equal to 250 at the beginning of therapy. Additionally, these results estimate that this increase in response odds could be as small as 27 percent, and as large as 141 percent.

Hence, the resulting estimates of the difference in proportions relative risk and odds ratios based on the same data will all agree in terms of the direction of association and the resulting confidence intervals will all agree with the inclusion or exclusion of the null value.

3 Hypothesis Testing - Understanding the general concept

- Using sample evidence, the hypothesis testing is used to assess the plausibility of the hypothesis.
- The test demonstrates the plausibility of the hypothesis and acknowledges the inconsistency between two comparable groups.
- Statistical analysts test the hypothesis by evaluating and calculating a random population sample.
- Two complementary methods of resolving uncertainty in sample-based comparisons when making assumptions about unknown population comparisons are confidence intervals and hypothesis testing.

➢ The general concept of Hypothesis testing

There are two conflicting truths sometimes referred to as the null hypothesis, and the alternative hypothesis.

- Null hypothesis (**H_0**)- The two populations do not differ.
- Alternative hypothesis (**H_A**) – The two populations vary.

Hypothesis testing starts by assuming that H_0 is truth.

If H_0 is truth then the estimated difference between samples is expected to be close to zero and within plus or minus two standard errors of zero again with large samples, using the result for the central limit theorem.

❑ Hypothesis Testing – The Paired Approach

The method of measuring the p-value is referred to as a **paired t-test** for a hypothesis test of mean difference between two populations.

- "paired" because of the study design.
- "t-test" because the sampling distribution is a t-distribution.

If two competing hypotheses are set up about the unknown population means, a third hypothesis can be formed assuming the difference between the other two.

Table 9 Competing Hypotheses for paired t-test. Source and credit. Coursera

Hypothesis 1	Hypothesis 2	Hypothesis 3
$\mathbf{H_0: \mu_1 = \mu_2}$	$H_0: \mu_1 - \mu_2 = 0$	$H_0: \mu_{diff} = 0$
$\mathbf{H_A: \mu_1 \neq \mu_2}$	$H_A: \mu_1 - \mu_2 \neq 0$	$H_A: \mu_{diff} \neq 0$

If H_0 is taken to be true: Assuming that the underlying population means are same and that their difference is zero, and then computing the difference in means i.e., how far is the observed mean from the expected difference (from 0) in terms of standard errors.

$$t = \frac{\bar{x}_{diff} - 0}{\widehat{SE}(\bar{x}_{diff})} = \frac{\bar{x}_{diff}}{\widehat{SE}(\bar{x}_{diff})}$$

After this the difference is translated into p-value, that further helps in taking the decision for hypothesis testing.

If the p-value is $< 0.05 \rightarrow$ In favor of alternative hypothesis (H_A), we reject the null hypothesis (H_0).

If the p-value is $\geq 0.05 \rightarrow$ We consider to accept the null hypothesis (H_0).

❑ Hypothesis Testing – The Unpaired Approach

The method of measuring the p-value is referred to as an **unpaired t-test** for a hypothesis test of mean difference between two independent populations.

- "unpaired" because of the study design.
- "t-test" because the sampling distribution is a t-distribution.

If two competing hypotheses are set up about the unknown population means, a third hypothesis can be formed assuming the difference between the other two.

Table 10. Competing Hypotheses for unpaired t-test. Source and credit. Coursera

Hypothesis 1	Hypothesis 2	Hypothesis 3
$\mathbf{H_0: \mu_1 = \mu_2}$	$H_0: \mu_1 - \mu_2 = 0$	$H_0: \mu_{diff} = 0$
$\mathbf{H_A: \mu_1 \neq \mu_2}$	$H_A: \mu_1 - \mu_2 \neq 0$	$H_A: \mu_{diff} \neq 0$

If H_0 is taken to be true: Assuming that the underlying population means are same and that their difference is zero, and then computing the difference in means i.e., how far is the observed mean from the expected difference (from 0) in terms of standard errors.

$$t = \frac{\bar{x}_{diff} - 0}{\widehat{SE}(\bar{x}_{diff})} = \frac{\bar{x}_{diff}}{\widehat{SE}(\bar{x}_{diff})} \quad \text{where, } \widehat{SE}(\bar{x}_{diff}) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

After this the difference is translated into p-value, that further helps in taking the decision for hypothesis testing.

If the p-value is $< 0.05 \rightarrow$ In favor of alternative hypothesis (H_A), we reject the null hypothesis (H_0).

If the p-value is $\geq 0.05 \rightarrow$ We consider to accept the null hypothesis (H_0).

❑ A note on p-value

- A p-value is a probability, a number between zero and one.
- Small p-values mean that if null is valid, the sample findings are unlikely.
- Alternatively, the p-value would suggest that the probability of obtaining a study results as extreme or more extreme, as unlikely or more unlikely than by likelihood of itself, if the null hypothesis were true.
- If the p-value is $< 0.05 \rightarrow$ In favor of alternative hypothesis (HA), we reject the null hypothesis (H0).

  If the p-value is $\geq 0.05 \rightarrow$ We consider to accept the null hypothesis (H0).
- The concept of two types of errors, namely Type 1 and Type 2 errors, occurs here.
- The confidence level is referred to as the probability to accept a true null hypothesis.
- The power of the test is the probability to reject a false null hypothesis. The significance level, sample size, alternative hypothesis, and nature of the test, play a great role in determining this.
- Type I error occurs when the true Null hypothesis is rejected. The probability of making Type I error is α (always known) and can be minimized by decreasing α (the level of significance).
- Type II error occurs when the false null hypothesis is not rejected. The probability of making Type II error is β (rarely known) and can be minimized by increasing difference of alternative hypothesis, increasing sample size, increasing α, or by choosing another test.

Table 11. Types of error. Source. Original.

		DECISION TAKEN IN HYPOTHESIS TESTING	
		ACCEPT	**REJECT**
THE TRUE STATE OF NULL HYPOTHESIS	**TRUE**	**Correct Decision** $(1 - \alpha)$ ☺	**TYPE I ERROR** (α)
	FALSE	**TYPE II ERROR** (β)	**Correct Decision** $(1 - \beta)$ ☺

❑ Hypothesis Testing – The z-test Approach

- The two-sample z-test presents a means for obtaining a p-value to test two competing hypotheses on the true proportions of the binary outcome among both populations as well as the competing hypotheses.

$$H_0: p_1 = p_2$$

$$H_A: p_1 \neq p_2$$

- These two hypotheses can be equivalently expressed in terms of Risk Difference (RD), Relative Risk (RR) and Odds Ratio (OR).

Table 12. hypotheses expressed in terms of Risk Difference (RD), Relative Risk (RR) and Odds Ratio (OR) for z-test. Source and credit. Coursera

Risk Difference	Relative Risk	Odds Ratio	$\ln(Relative\ Risk)$	$\ln(Odds\ Ratio)$
H_0: RD = 0	H_0: RR = 0	H_0: OR = 0	$H_0: \ln(RR) = 0$	$H_0: \ln(OR) = 0$
H_A: RD ≠ 0	H_A: RR ≠ 0	H_A: OR ≠ 0	$H_A: \ln(RR) \neq 0$	$H_A: \ln(OR) \neq 0$

- As such only one p-value (and hence, one hypothesis test) is needed for all measures of association comparing binary outcomes between two populations.
- The test is performed using observed difference in proportions (risk difference: $\widehat{p_1} - \widehat{p_2}$) and its estimated standard error ($\widehat{SE}(\widehat{p_1} - \widehat{p_2})$).
- The decision for hypothesis testing is taken based on p-value (translated from the difference), as mentioned earilier.

$$Z = \frac{(\widehat{p_1} - \widehat{p_2}) - 0}{\widehat{SE}(\widehat{p_1} - \widehat{p_2})}$$

❑ Hypothesis Testing – The Chi-square test Approach

- The chi-square test is performed for comparing binary (or categorical) results across two populations or more.
- The results of the chi-square test and the two-sample z- test are similar in a particular case of two proportions being measured between two populations – both rely on the same Central Limit Theorem based outcome.
- For one test the chi-square test could be applied to further comparisons.
- The two competing hypotheses (observed and expected):

$$H_O: p_O = p_E$$

$$H_A: p_O \neq p_E$$

- These two hypotheses can be equivalently expressed in terms of Risk Difference (RD), Relative Risk (RR) and Odds Ratio (OR).

Table 13 Hypotheses expressed in terms of Risk Difference (RD), Relative Risk (RR) and Odds Ratio (OR) for chi-square test. Source and credit. Coursera

Risk Difference	Relative Risk	Odds Ratio	$\ln(Relative\ Risk)$	$\ln(Odds\ Ratio)$
H_0: RD = 0	H_0: RR = 0	H_0: OR = 0	$H_0: \ln(RR) = 0$	$H_0: \ln(OR) = 0$
H_A: RD ≠ 0	H_A: RR ≠ 0	H_A: OR ≠ 0	$H_A: \ln(RR) \neq 0$	$H_A: \ln(OR) \neq 0$

- The measure of difference of the observed value from the expected, in chi-square test is;

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

- The decision for hypothesis testing is taken based on p-value (translated from the difference), as mentioned earilier.

❑ Hypothesis Testing – The Fischer's Exact test Approach

- Another way to achieving a p-value is called **Fisher 's exact test** when comparing the proportions of two populations.
- This can be used regardless of sample size (which requires large sample sizes for both the two-sample z and Chi-Square test).
- Results for the p-value from this test would be somewhat close in values of the two-sample or Chi-Square test, but in a small sample case, probably again the values would tend to differ.
- But it gets computationally intensive in cases of large sample sizes.
- The hypothesis testing procedure is similar to that of chi- square test, but it differs in the calculation for p-value (can be understood by an example).
- Example: Let us consider a table with literate and illiterate males and females.

Table 14. A table of literate and illiterate males and females. Source. Original.

	Males	Females	Total
Literate	p	q	p+q
Illiterate	r	s	r+s
Total	p+r	q+s	p+q+r+s = n

$$P = \frac{\binom{p+q}{p}\binom{r+s}{r}}{\binom{n}{p+r}} = \frac{\binom{p+q}{q}\binom{r+s}{s}}{\binom{n}{q+s}} = \frac{(p+q)!\ (r+s)!\ (p+r)!\ (q+s)!}{p!\ q!\ r!\ s!\ n!}$$

5 Hypothesis Testing - Comparing more than two populations with one test

- The hypothesis test we have considered for comparing two populations via two samples can be extended to allow for the comparison of more than two populations in one test.
- The general approach to hypothesis testing is the same as with two populations, but to competing hypotheses we have were as follows;
- The null hypothesis (H_O): All population parameters (means, proportions, incidence rates) are the same.
- The alternative hypothesis (H_A): At least two of the populations vary in mean values or proportions or incidence rates.
- Statistically significant result for such a test results indicates that there are population differences, but does not specifically identify the populations that are different.
- The names of the tests comparing multiple populations are as follows;
    - ❖ For means – Analysis of variance (ANOVA) extends the two samples unpaired t-test.
    - ❖ For proportions – The chi-square test can be extended.
    - ❖ For incidence rates/ survival curves for individual time- to -event outcomes: The log-rank test can be extended.

❑ Hypothesis Testing – The Analysis of Variance (ANOVA)

- To interpret a p-value from a hypothesis test testing for any mean differences between more than two populations, the method used, is called Analysis of Variance or ANOVA.
- The competing hypotheses are;

  $H_{0:}\ \mu_1 = \mu_2 = \ldots\ldots\ldots\ldots\ldots\ \mu_k$

  H_A: μs vary for at least two populations.
- Yet another strategy has to shape a lot of two sample t-tests for each potential comparison between two groups.
- Presuming the null hypothesis is valid, and measuring to some degree of difference in what was found in the sample as compared with the predicted sample means being under null hypothesis.
- ANOVA, this indicator of discrepancy called the F-statistic., that is compared to an F-distribution for obtaining a p-value.
- The measurement for the F-statistic can be done using one-way ANOVA or the regression analysis.
- After this the difference is translated into p-value, that further helps in taking the decision for hypothesis testing.
- If the p-value is $< 0.05 \rightarrow$ we reject the null hypothesis (H_0), i.e., at least means for two populations vary.
- If the p-value is $\geq 0.05 \rightarrow$ we accept the null hypothesis (H_0).

❑ Hypothesis Testing – Chi-Square test for comparing proportions

- Chi-squared tests can be used to test differences between two or more populations with one test for population level.
- The method of getting a p-value is an extension of the chi-square test mentioned earlier.
- The two competing hypotheses;

  $H_{0:}\ p_1 = p_2 = \ldots\ldots\ldots\ldots\ldots\ p_k$

  H_A: ps vary for at least two populations.
- Degrees of freedom for this test depend on how many populations are being compared.
- A significant p value when that less than 0.05 means that at least two populations have different underlying proportions. This result does not identify the specific groups that differ.

❑ Hypothesis Testing – The log rank test approach

- The method of interpreting a p-value for the hypothesis test comparing survival curves (S(t)) and hence, incidence rates between more than two populations, is referred to as log rank test.
- The two competing hypotheses;

  H0: $S_1(t) = S_2(t) = \ldots\ldots\ldots\ldots\ldots\ S_k(t)$

  HA: S(t)s varies for at least two populations.
- The log-rank test is popular for testing differences in survival curves at population level and hence, the underlying incidence rates between two or more populations for time-to-event data collected at the individual level.
- For this a significant "p-value", something lesser than 0.05, means that from a statistic point of view, at least two populations have different underlying population level survival curves.
- But this test alone does not identify the specific populations that differ, a comparison for each two group is done to find where those differences are.

Abbreviations

PPT	PowerPoint presentations
ANOVA	Analysis of Variance
S.D.	Standard deviation
S.E.	Standard Error
C.I.	Confidence Interval
CD4	Cluster of Differentiation 4
HIV	Human Immunodeficiency Virus
RR	Relative Risk
OR	Odds Ratio
RD	Risk Difference
Diff.	Difference

References

1. Online Courses & Credentials from Top Educators. Join for Free. (n.d.). Retrieved from https://www.coursera.org/learn/hypothesis-testing-public-health/home/info
2. The Sampling Distribution and the Central Limit Theorem. (n.d.). Retrieved from https://www.ultracad.com/sampling_distribution.htm
3. Joseph, N. (2020, February 19). Calculation and Interpretation of Confidence Intervals. Retrieved from https://analystprep.com/cfa-level-1-exam/quantitative-methods/confidence-intervals/