META-ANALYSIS ON THE EFFECTS OF COVID-19 ON THYROID DISORDERS

by

İpek Dağdeviren

Submitted to Graduate School of Natural and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in
Data Science

Yeditepe University
2022

META-ANALYSIS ON THE EFFECTS OF COVID-19 ON THYROID DISORDERS

APPROVED BY:

Assist. Prof. Dr. Elif Çiğdem Keleş ……
(Thesis Supervisor)
(Yeditepe University)

Prof. Dr. Gürhan Küçük ……
(Yeditepe University)

Prof. Dr. Dilek Altaş Karaca ……
(Marmara University)

DATE OF APPROVAL: ..../..../20....

I hereby declare that this thesis is my own work and that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I have fully cited and referenced all material and results as required by these rules and conduct, and this thesis study does not contain any plagiarism. If any material used in the thesis requires copyright, the necessary permissions have been obtained. No material from this thesis has been used for the award of another degree.

I accept all kinds of legal liability that may arise in case contrary to these situations.

Name, Last name İpek Dağdeviren

Signature ……....

ACKNOWLEDGEMENTS

It is with immense gratitude that I acknowledge the support and help of my advisor Assist. Prof. Dr Elif Çiğdem Keleş and Assist. Prof. Meliha Melin Uygur, MD for the medical consultancy. Pursuing my thesis under her supervision has been an experience which broadens the mind and presents an unlimited source of learning.

ABSTRACT

META-ANALYSIS ON THE EFFECTS OF COVID-19 ON THYROID DISORDERS

Millions of people were affected both by COVID-19 and thyroid diseases across the world but previous works are limited to identify the effects of COVID-19 on thyroid disorders. Thyroid holds a critical role to regulate numerous metabolic processes throughout the body, a few changes in it, may have affect on other systems on the human body. During the pandemic, thyroid disorders get increased and those disorders are usually interpreted as autoimmune mechanisms. Present meta-analyses are completed with limited parameters and just focus on a particular topic which prevents to act in a holistic approach. Our study focuses on the relation between thyroid function test measures, thyroid dysfunction, comorbidity and COVID-19 severity; also thyroid function test measures and mortality caused by COVID-19; thyroid function test measures and COVID-19 existence are observed. In total, 16 meta-analyses are completed and aim to improve the evidence level of the studies related to thyroid disorders and COVID-19 which will contribute to the medical sciences. 4260 studies were reviewed from PubMed and Science Direct search engines, and after applying the exclusion criteria, 13 studies with a sample group of 2829 individuals were included in the meta-analyses. PRISMA 2020 Checklist flows were tracked, for statistical calculation by using the RevMan and MedCalc software. All in all, thyroid dysfunction patients tend to undergo more severe COVID-19. The relation between thyroid disorders and COVID-19 was expressed with meta-analyses by using many parameters, aiming to increase the reliability level of the related studies.

ÖZET

COVID-19'UN TİROİT BOZUKLUKLARI ÜZERİNDEKİ ETKİLERİ ÜZERİNE META-ANALİZ

Dünya genelinde COVID-19 ve tiroit hastalıklarından milyonlarca insanın etkilendiği bilinmektedir. Literatürdeki çalışmalar COVID-19'un tiroit üzerindeki etkilerini açıklamakta yetersiz kalmaktadır. Tiroit, insan vücudundaki birçok metabolik süreci düzenlemede kritik bir role sahip olduğundan, işleyişindeki bir değişikliğin etkisi vücuttaki farklı sistemlere de olacaktır. COVID-19 pandemisi ile artan tiroit bozuklukları genellikle otoimmün mekanizmalara bağlı olarak yorumlanmaktadır. Mevcut meta-analizler sadece belirli parametrelere odaklandığından bütüncül bir yaklaşım sergilememektedir. Bu çalışma tiroit fonksiyon test değerleri, tiroit bozuklukları, komorbidite ve COVID-19 ciddiyeti; tiroit fonksiyon test değerleri ve COVID-19 kaynaklı ölüm vakaları, tiroit fonksiyon test değerleri ve COVID-19 varlığı parametrelerine odaklanmıştır. Toplamda 16 meta-analiz yapılmıştır. Tıp bilimlerine katkı sağlamanın yanında, COVID-19'un tiroit bozukları üzerine etkisi ile ilgili çalışmaların da güncel literatürler ile birlikte kanıt düzeyini iyileştirmeyi amaçlanmaktadır. Literatür taramasında PubMed ve Science Direct arama motorları kullanılmış olup toplamda 4260 makale taranmıştır. Dışlama kriterleri sonrasında toplamda 2829 birey ile 13 çalışma meta-analize dahil edilmiştir. İstatistik hesaplamalar yapılırken PRISMA 2020 kontrol listesi takip edilerek, RevMan ve MadCalc yazılımlarından yararlanılmıştır. Yapılan meta-analizler ile tiroit bozukluğu olan bireylerin COVID-19 seyirlerinin daha ciddi olduğu sonucuna varılmıştır. Bu çalışma ile bir çok parametre kullanarak tiroit bozuklukları ve covid19 ilişkisi meta analizler ile ortaya konularak konu hakkındaki çalışmaların kanıt seviyesi artırılmaya çalışılmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

df	Degree of freedom
OR	Odds ratio
CI	Confidence interval
T4	Thyroxine
T3	Triiodothyronine
TBG	Thyroxine-binding globulin
TSH	Thyroid-stimulating hormone
SARS	Severe acute respiratory syndrome
MERS	Middle East respiratory syndrome
FDA	U.S. Food and Drug Administration
CD14+	Cluster of differentiation 14
IL-6	Interleukin 6
TNF-α	Tumor necrosis factor
ACE2	Angiotensin-converting enzyme 2
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
TFT	Thyroid function test
IQR	Inter quartile range
ICU	Intensive care unit
SD	Standard deviation
NTIS	Non-thyroidal illness syndrome
ES	Effect size
SMD	Standardized mean difference

1. INTRODUCTION

Meta-analysis is a statistical technique, aimed to combine the findings of more than one independent study conducted on a particular topic and having statistical analysis obtained from the outputs. It is a process to interpret the variety among the outputs and to come up with more reliable and accurate outputs. Also, the significance level of studies gets increases. [1]

1.1. OBJECTIVES OF META-ANALYSIS

There exist severel objectives to conduct a meta-analysis. Some of them served below [2, 3, 4]:

- Due to meta-analysis combines the multiple studies, it expands the sample size of studies that have been completed with small samples groups. As a result, it increases the precision of the parameters. Also, the statistical significance level gets increases.
- It evaluates and observes the reasons of contradictoriness generated from the scientific literature.
- To discover the source of heterogeneity emerges among the studies.
- To examine the changes that are not considered in the primary studies but are assumed to have an effect.
- Help to shape the future works.
- According to the findings of the meta-analysis, new research topics can emerge for future works.
- Finding answers to questions that are not considered at the beginning of the study.
- Increasing the effect size predictions.
- To come up with decisions in terms of uncertainty when a contradiction occurs among the outputs of independent studies.

1.2. HISTORY OF META-ANALYSIS

In 1904, one of the first meta-analysis was completed by Karl Pearson which focuses on the effects of inoculation and typhoid fever. He used five datasets which are consisting of 2x2 tables and they include the infection rate of vaccinated and unvaccinated sample groups. Then, from each table, he calculated the correlation between vaccination and infection rate. Finally, in order to summarize the mode of the relationship, he calculated the mean value of those correlations and used that value as the effect size of the treatment. After comparing the value that he obtained with the effect size of the smallpox vaccine, he came to conclusion that: vaccination for typhoid fever is ineffectual. Pearsons' study contains all the features that are needed for a successful meta-analysis. [5]

Until 1970's meta-analysis generally applied on health sciences. After 1970's due to accessing to studies becoming easier, meta-analysis get popular [4]. Today, meta-analysis can also be conducted on social, medical, education, phycology and biomedical sciences as well. [5]

1.3. ADVANTAGES OF META-ANALYSIS

The amount of scientific research is increasing rapidly day by day. Independent studies which are conducted on the same particular subject usually result in different conclusions. To interpret these results, more reliable and more inclusive methods are needed. Meta-analysis is one of the best technique to achieve this.

- It is a time and energy-saving method because researchers complete a literature review on just one particular topic.
- Helps to clarify the questions and the possible solutions to answer the questions.
- Serves independent studies in a holistic way and outputs obtained from studies can be generalized.
- Helps to get rid of the inconsistency issues among different studies.
- Serves more systematic and well-planned approach, due to there exist some mandatory steps to complete a meta-analysis.

According to Abramson (1994), the advantages of combining the multiple studies which have been done in the same topic [6]:

- If the individual studies have similar findings, it will increase the reliability of each study.
- If the individual studies have been completed with small sample groups, it is difficult to classify them as statistically significant. Since meta-analysis combines the findings the multiple studies it is able to cope with that problem.
- If individual studies resulted in different findings, meta-analysis tries to explain the causes of the problems that make up this difference. As a result, meta-analysis helps to explore new hypotheses and new findings.

1.4. DISADVANTAGES OF META-ANALYSIS

Despite the advantages of meta-analysis, on the other hand, there exist some disadvantages of meta-analysis as well. [2, 3, 4]

- According to Glass'(2015) apples-oranges problem, he evaluated that oranges and apples are compared in meta-analysis approach. He argued that it is not feasible to come reasonable generalizations when applying different measuring techniques, variable definitions and studying on not very similar topics. He emphasized that there is no need to compare studies that are identical in all respects because they will give very similar results. Just different studies should be compared or synthesized. [7]
- Since there exist many publications related with the same research, this situation leads to bias and reduces the reliability of meta-analysis.
- When performing a meta-analysis, using not similar studies or inclusion of erroneous studies into the meta-analysis is possible.
- Bias in publications and including poor studies into the meta-analysis is another pain point.
- Combining/evaluating successful and poor studies into the same meta-analysis leads to criticism.

- If studies are reviewed on a particular topic, it is examined that the majority of the studies concluded as different results in terms of statistically. So, bias predictions take place in meta-analysis.

1.5. STEPS OF META-ANALYSIS

There are eight mandatory steps to perform a meta-analysis. [5]

1. Identifying the problem
2. Determining the inclusion and exclusion criterias
3. Completing the literature review according to the inclusion criteria
4. Extracting, clustering and coding the studies in terms of methodology and outputs
5. Applying statistical techniques
6. Combining the outputs of the studies
7. Establishing the relationship between the outputs of the studies and characteristics of meta-analysis.
8. Re-reviewing, interpreting and presenting the results

1.6. HETEROGENEITY IN META-ANALYSIS

Each study included into the meta-analysis ended with different findings/effects size. It is an expected case and it does not matter whether having difficulties or not, the key point is: All those difficulties can be either ignored or not? In meta-analysis before including combined studies into it, firstly heterogeneity among the studies should be reviewed. In order to do that there exist statistical and graphical methods. [8]

If the confidence interval of the study' outputs has minor intersect or have not intersect, this case generally indicates that having heterogeneity and forest plot chart should be reviewed. For determining the heterogeneity, the Q-statistics test is the most widely and most commonly used one. Unfortunately, it is not capable to explain the factors causing the heterogeneity.

Usually, the cause of heterogeneity among the studies is unexplainable, the random effects model should be preferred. Explicitly, if heterogeneity exists random effects model if

heterogeneity does not exist fixed effect model should be used. After applying the random effects model, in order to explain the cause of heterogeneity ANOVA or meta-regression methods can be used. [9]

Some factors causes of heterogeneity are different working patterns, quality of work, bias arising from the choice of study, other bias effects, usage of different statistical methods. [9]

1.6.1. Testing the Heterogeneity

Q-Statistics

Q-statistics heterogeneity test is offered by Cochran in 1954. Known as Cochran's Q statistics. According to West, Gartlehnern and Mansfield(2010), Cochran's Q test provides a method for testing for differences between three or more matched sets of frequencies or proportions. Based on a chi-square distribution, it generates a probability that, when large, indicates larger variation across studies rather than within-subjects within a study. A limitation of Cochran's Q test is that it might be underpowered when few studies have been included or when event rates are low. Therefore, it is often recommended to adopt a p-value (rather than 0.05) as a threshold for statistical significance when using Cochran's Q test to determine statistical heterogeneity. [10] The formula of Q-value is declered in equation 1.1. [11]

$$Q = \sum_{i=1}^{k} w_i(\theta_i - \theta)^2 \quad \text{where } \theta \text{ is } \theta = \frac{\sum_{i=1}^{k} w_i\theta_i}{\sum_{i=0}^{k} w_i} \tag{1.1}$$

Q-statistics is just useful to evaluate whether heterogeneity exists or not, it does not give any information about the size of heterogeneity. By applying with chi-square distribution, heterogeneity can be tested in meta-analysis studies. [12]

1.6.2. Heterogeneity Measures

Tau(τ)^2 Statistics

τ^2 statistics is one of the way to determine the amount of heterogeneity. It describes the variance among the studies and obtained from Cochran's Q statistics. Also, do not effect

by the sample size and amount of study included in the meta-analysis. There are different methods to predict τ^2 in meta-analysis such as Der-Simoniasn Laird method, Hedges and Olkin method, Malzahn and Böhning method etc. The most commonly used method is Der-Simoniasn Laird. [11] Formula 1.2 shows the calculation of τ^2. [11]

$$\tau^2 = \frac{(Q - df)}{\sum_{i=1}^{k} W_i - \frac{\sum_{i=1}^{k} W_i^2}{\sum_{i=1}^{k} W_i}} \tag{1.2}$$

Then the variance among the studies will be like the following, k determines the number of studies included into the meta-analysis, formula 1.3 can be applied. [11]

$$\tau^2 = \begin{cases} \dfrac{Q - (k-1)}{\sum_{i=1}^{k} W_i - \frac{\sum_{i=1}^{k} W_i^2}{\sum_{i=1}^{k} W_i}}, & Q > (k-1) \\ 0, & Q \leq (k-1) \end{cases} \tag{1.3}$$

Since the value of τ^2 is not fit to the normal distribution, in order to determine 95 percent confidence interval (CI) limits, formula 1.4 can be applied. [11]

$$CI = e^{0.5 * \ln\left(\frac{Q}{k}\right) \pm 1.96 * B} \tag{1.4}$$

If Q-value is greater than the number of studies included into the meta-analysis (Q > k), B value can be found by the formula 1.5. [11]

$$B = 0.5 * \frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{(2k-3)}} \tag{1.5}$$

If Q-value is equal to or less than the number of studies included into the meta-analysis (Q ≤ k) B value can be found by the formula 1.6. [11]

$$B = \sqrt{\frac{1}{2(k-2)\left[1 - \left(\frac{1}{3(k-2)^2}\right)\right]}} \tag{1.6}$$

If τ^2 is equal to zero, predicted variance among the studies can be ignored, in conclusion, studies are homogeneous. In this case, the results are same with fixed effect model. If τ^2 is greater than zero, predicted variance among the studies can not be ignored, in conclusion, studies are not homogeneous. [13]

H Statistics

H statistics is firstly offered by Higgins ve Thompson in 2002. Also, it is known as Birge's ratio. It is obtained from Cochran's Q statistics and can be found by the formula 1.7 [13]:

$$H^2 = \begin{cases} \frac{Q}{k-1}, & Q > df \\ 1, & Q \leq df \end{cases} \quad (1.7)$$

H^2 value can be taken value between 1 and infinity. If the H value is equal to one, this indicates that well homogeneity. 95 percent confidence interval can be found by the formula 1.8 and evaluated by the formulas 1.9 and 1.10 [13]:

$$\exp[\ln(H) \pm 1.96 * SE(\ln(H))] \quad (1.8)$$

Standard error for ln(H), if Q > k:

$$SE(\ln(H)) = 0.5 * \left(\frac{\ln(Q) - \ln(k-1)}{\sqrt{2Q} - \sqrt{(2k-3)}}\right) \quad (1.9)$$

Standard error for ln(H), if Q ≤ k:

$$SE(\ln(H)) = \sqrt{\frac{1}{2(k-2)}\left(1 - \frac{1}{3(k-2)^2}\right)} \quad (1.10)$$

H value will cope with bias when less than eight studies are included into the meta-analysis. The confidence interval of the H value should contain one value, otherwise, heterogeneity will exist. Since the number of studies is less, the stability of the H value is less. As a result, detecting the heterogeneity gets difficult. As the number of studies increases, also the stability of the H value gets increases.

I^2 Statistics

I^2 shows the ratio of the total variance related to the effect size and is used to measure the heterogeneity in meta-analysis. Conversely to Cochran's Q, the amount of studies does not effect the I^2. It takes the values between zero and one and is used to express heterogeneity in terms of percentage. If I^2 is less than 25%, this inferred us low heterogeneity, if it is between 25% and 50%, this inferred us moderate heterogeneity and over 50-75% inferred us high heterogeneity [10]. I^2 value can be found by the following formula 1.11 where df = degrees of freedom [11]

$$I^2 = \left(\frac{Q - df}{Q}\right) * 100\% \quad (1.11)$$

Comparing the Measures of Heterogeneity

Q, p, τ^2, H and I^2 are the ways to measure heterogeneity. According to Borenstein (2011), since all the indices are based on Q (in relation to df), it follows that all will be low (or zero) if the total dispersion is low relative to the error within studies, and higher if the total dispersion is high, relative to the error within studies. However, the various measures of heterogeneity build on this core in different ways which make each useful for a specific purpose. [11] Table 1.1 serves the factors effecting measures of dispersion for heterogeneity measures.

Table 1.1. Factors effecting measures of dispersion

	Range of Possible Values	Depends on the Number of Studies	Depends on Scale
Q	$0 \leq Q$	X	
p	$0 \leq p < 1$	X	
τ^2	$0 \leq \tau^2$		X
H	$1 \leq H < \infty$	X	
I^2	$0\% \leq I^2 < 100\%$		

1.7. STATISTICAL MODELS FOR META-ANALYSIS

After determining the studies which will be included into the meta-analysis, then, outputs of the studies should be combined by using the statistical models. There exist two types of statistical models for meta-analysis. Those are “fixed effects model” and “random effects model”. Homogeneity and heterogeneity of the studies lead to determining the type of model. [4] According to Spineli and Pandis (2020), the choice of the appropriate model for the analysis is critical to ensure the credibility of the results and depends on both the goals of the analysis and the assumptions of the models. [14]

In both models, the p-value is one of the fundamental terms and is used to determine whether any explainable difference occurs or not between two variables or two groups. If the p-value is less than 0.05, it indicates that there exists an explainable difference between the variables or groups. In meta-analysis either having explainable difference or not among the individual studies, it is interpreted by using the p-value. Less than 0.05 p-value

demonstrates that there are explainable difference that takes place and meta-analysis study is heterogeneous. [2]

1.7.1. Fixed Effects Model

The fixed effects model assumes that all studies used for the meta-analysis share a common effect size [11]. One of the limitations to combining those studies is: samples sizes of them may be different from each other. This situation may indicate that also the universes of the studies may be different. Fixed model supposes that studies that are included into the meta-analysis have the same universe sizes and as a result, standard deviations of them are equal to zero. If the universe sizes of the studies are not equal, random effects model should be used and different standard deviations values from zero should be assumed. [2] If the result of the heterogeneity test for individual studies is low, then the fixed effects model can apply. [4]

1.7.2. Random Effects Model

Random effects model assumes that all studies which are included into meta-analysis have different effect sizes. Because each study has been completed by different researchers in different environments. So, they are heterogeneous. It is stated that the universe sizes of the studies vary and their standard deviations are not equal to zero. The variance value among the studies may be greater or lower than the expected variance value. If the expected variance value is lower, most probably both fixed effects model and random effects model will serve the same output. [4]

If the findings of the studies are desired to generalize fixed effects model should not be prefered. Random effects model based on the idea: The actual effect size may vary among the studies according to the age of attendees, education level, sample size etc. Thus, the actual effect size is distributed according to an average. [2]

1.8. TERMINOLOGY IN META-ANALYSIS

Effect size: "It is a statistical concept that measures the strength of the relationship between two variables on a numeric scale ("Effect Size")." For example, if we weighted a hundred women and men, then calculate the mean value for each gender. We will find that men are fatter than women. Effect size is the difference between the weight of the men and the weight of the women. If the difference between men and women weights' are larger, also the effect size will be high. "Statistic effect size helps us in determining if the difference is real or if it is due to a change of factors ("Effect Size")." Effect size is related with multiple studies and combines them into a single analysis in meta-analysis. In studies, effect size is used to determine how much the dependent variable has effect on the independent variable. [15]

Effect size can be measured by using standardized mean difference, odds ratio and correlation coefficient. The standardized mean difference is used when data are calculated according to the mean value and standard deviation of the two groups. Odds ratio effect size is used when data are calculated according to the results of binary groups such as either having an event or not. Correlation coefficient effect size is used when data are calculated by using the correlation value between two variables. [16]

If studies included into the meta-analysis uses different scales, combining the mean standardized difference does not make any sense. In that case, effect size of each independent study should be standardized. In order to do that standardized mean difference is used as the effect size (Cohen's d or Hedges's g). Cohen's d and Hedges's g are the coefficients that are used in meta-analysis to calculate the effect size. If sample sizes are different or smaller than 20, then Hedges g gives more accurate values. [17] Despite, their calculation formulas being different, their results are similar to each other. In small sample groups, large Cohen's d value causes bias. To eliminate this bias, Hedge's changes Cohen's formula and then defines fair effect size which calls as Hedge's g. [16] Formula 1.12 and 1.13 for both Cohen's d and Hedge's g are shown below:

$$\text{Cohen's d: } d = \frac{\bar{x}_{test} - \bar{x}_{control}}{S_{pooled}} \quad S_{pooled} = \sqrt{\frac{(n_{test} - 1) * s_{test}^2 + (n_{control} - 1) * s_{control}^2}{(n_{test} + n_{control})}} \tag{1.12}$$

$$\text{Hedge's g: } d = \frac{\bar{x}_{test} - \bar{x}_{control}}{S_{pooled}} \quad S_{pooled} = \sqrt{\frac{(n_{test} - 1) * s_{test}^2 + (n_{control} - 1) * s_{control}^2}{(n_{test} + n_{control} - 2)}} \tag{1.13}$$

d = Effect size according to Cohen's d

n = sample size

s = standart deviation

S_{pooled} = total standart deviation

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 evaluated as small relative size, if the effect size is between 0,5 and 0,8 evaluated as medium relative size and if the effect size is greater or equal than 0,8 evaluated as large relative size. (table 1.2) [18]

Table 1.2. Standardized mean difference relative size

Relative Size	**Effect Size**
Small	0,2 ≤ effect size < 0,5
Medium	0,5 ≤ effect size < 0,8
Large	0,8 ≤ effect size

For correlation coefficient difference if the effect size is between 0,1 and 0,25 evaluated as small relative size, if the effect size is between 0,25 and 0,4 evaluated as medium relative size and if the effect size is greater or equal than 0,4 evaluated as large relative size. (table 1.3)

Table 1.3. Correlation coefficient relative size

Relative Size	**Effect Size**
Small	0,1 ≤ effect size < 0,25
Medium	0,25 ≤ effect size < 0,4
Large	0,4 ≤ effect size

Odds

It indicates the relationship between the probability of an event and the probability of non-event. Calculated by dividing the probability of an event by the probability of non-event. For example, when rolling a fair dice probability of having one is 1/6 and not having one is 5/6, so the odds value becomes (1/6)/(5/6) = 1/5.

Odds ratio (OR)

It can be obtained by proportioning two different odds values. Defines the relationship between exposure and outcome. “The odds ratio represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of the exposure. ("Odds Ratio")” It can take values between zero and infinity. If greater than one this means the increased occurrence of an event. If less than one this means the decreased occurrence of an event. [19] An example of calculating the odds ratio is shown in table 1.4 and formula 1.14:

Table 1.4. Odds ratio calculation

	Case Group	Control Group
Exposed	A	B
Unexposed	C	D

$$\text{Odds Ratio} = \frac{Odds\ that\ a\ case\ was\ exposed\ (A/C)}{Odds\ that\ a\ control\ was\ exposed\ (B/D)} = \frac{A * D}{B * C} \tag{1.14}$$

Relative Risk Ratio (RR)

It is a risk measure used in studies. Healthy individuals who have similar characteristics are divided into two groups: exposed to a factor and absence of exposure to a factor. Then follow-up them until the factor to cause the disease. At the end of the follow-up period, the incidence rate in the exposure group was proportioned to the absence of the exposure group. The obtained value is the relative risk ratio. If it equals one then, the risk of having an event is the same in both groups. If greater than one, risk of having an event in the exposure group is greater than having in absence of exposure. If less than one, risk of having an event in the exposure group is less than having in absence of exposure. [20] An example of calculating relative risk ratio is shown in table 1.5. and formula 1.15:

Table 1.5. Relative risc ratio calculation

	Case Group	Control Group
Exposed	A	B
Unexposed	C	D

$$\text{Relative Risc Ratio} = \frac{\frac{A}{A+B}}{\frac{C}{C+D}} \quad (1.15)$$

Correlation

Correlation is a statistical term, describes the strength of the relationship among two variables and measured by the correlation coefficient. Most widely used measure is Pearson's correlation coefficient (r) also known as Pearson product-moment correlation. It may take values between -1 to 1. Positive values describes if a variable increases, the other variable tends to increase as well or if a variable decreases, the other variable tends to decrease as well. Negative values describes if a variable is increases, the other variable tends to decrease or if a variable is decreases, the other variables tends to increase. Values near to -1 or 1 means strong linear association and value near to zero means low association. Pearson's correlation coefficient is only able to interpret when linear association take place among two variables. If variables are correleated with a more complex relationship, it may interpret as no correlation. [21]

Correlation coefficients values between ±1 and ±0.9 evaluated as very high correlation, between ±0.9 and ±0.7 evaluated as high correlation, between ±0.7 and ±0.5 evaluated as moderate correlation, between ±0.5 and ±0.3 evaluated as low correlation. [22]

Formula 1.16 is used to calculate Pearson's correlation coefficient (r), where n is the number of observations, X and Y are the variables.

$$\boldsymbol{r} = \frac{\boldsymbol{n}*(\sum(\boldsymbol{X},\boldsymbol{Y}) - (\sum(\boldsymbol{X})*\sum(\boldsymbol{Y})))}{\sqrt{\boldsymbol{n}*\sum(\boldsymbol{X^2}) - \sum(\boldsymbol{X})^2)*(\boldsymbol{n}*\sum(\boldsymbol{Y^2}) - \sum(\boldsymbol{Y})^2)}} \quad (1.16)$$

1.9. PUBLICATION BIAS IN META-ANALYSIS

Beyond the heterogeneity, bias is another serious problem that emerges in meta-analysis. In order to put forward a successful meta-analysis, clarifying no bias in the publication is one of the key points. One of the reasons behind this problem comes from combining the independent studies. Including low success or including/excluding some special studies into the meta-analysis, resulting in bias and error in outputs of the meta-analysis. Another reason is using multiple findings obtained from the same study. Moreover to meta-analysis, bias can also take place in individual studies as well. [23] Shapiro states that to cope with

bias in non-experimental studies some researchers assign quality weight to each study. [24, 25] According to Petitti, totally eliminating the bias is not feasible but it is great to identify variety and inconsistency among the studies [25].

The accessibility of the research outputs' does not only depend that either study have been published or not, also depends on where, when and how they were created. The proposed methods include extensive research for related studies also graphical and statistical methods to prevent publication bias. [26] For demonstrating the publication bias in meta-analysis: the funnel plot, rank correlation test and linear regression test are executed widely. Also, Rosenthal's FSN, Orwin's FSN and Duval and Tweedie's trim fill methods may apply. [13]

1.9.1. Significance Methods for Publication Bias in Meta-Analysis

Funnel Graph

A funnel graph is a scatter plot that describes the predicted effect size in each study versus each unit of the sample size of the studies. Mostly, the x-axis determines effect size and the y-axis determines the standard error, precision, sample size or variance. Outputs obtained from studies that have been conducted with a small sample size, the data will widely spread at the base of the graph. Despite that, outputs obtained from studies which have been conducted with a large sample size, the data will tightly spread at the top of the graph. All in all, if the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Bias is not the only reason for asymmetry in the graph. [2]

Rank Correlation Test (Begg Test)

Rank correlation test was defined by Begg ve Mazumdar in 1994. Also known as the Begg test. Publication bias causes a relationship between predicted effect size and standard errors of those predicted effect size. It observes the relation between standardized effect and its variances based on Kendal tau. Since rank correlation test is not powerful, it does not guarantee to identify publication bias. For k-times study standardized effect size can be found by the formula 1.17, 1.18, 1.19, 1.20, 1.21: [13, 27]

$$T_i^* = \frac{T_i - \bar{T}_.}{\sqrt{(\widehat{var}_i)}} \tag{1.17}$$

$$\bar{T}_{.} = \frac{\sum_{i=1}^{k} var_i^{-1} T_i}{\sum_{i=1}^{k} var_i^{-1}} \quad (1.18)$$

T_i and v_i shows the i^{th} study's effect size and sample variance, $\tilde{v}_i$ shows the variance of standardized effect size for i^{th} study:

$$\widetilde{var}_i = Var(T_i - \bar{T}_{.}) = var_i - (\sum_{i=1}^{k} var_i^{-1})^{-1} \quad (1.19)$$

$$z = \frac{t}{\sqrt{var(t)}} \quad (1.20)$$

t: Kendall' tau coefficient

var(t): sample variance of Kendall' tau coefficient

z: the value of standard normal distribution

$$Z = \frac{P - Q}{\sqrt{\frac{k(k-1)(2k+5)}{18}}} \quad (1.21)$$

P: number of the sorted couples in the same rank

Q: number of the sorted couples in the opposite rank

k: number of studies

Then, the value obtained should be compared with the standard normal distribution table value. [13, 27]

Linear Regression (Egger) Test

Linear regression test was defined by Egger in 1997. Also known as the Egger test. Aimed to test asymmetry in funnel graph. A regression line equation is produced. Independent variable is emerged by taking the inverse of studies' standard deviation overview over multiplication. The dependent variable has emerged from Egger' standardized effect size. If the intersection point of the regression line and the y-axis is near zero, bias does not exist. The regression equation that obtains from least square method is Y = a + bX. "a" variable is for asymmetry measure and if it is equal to zero then, there exists no bias. [13]

1.10. GRAPHICAL METHODS IN META-ANALYSIS

While observing the heterogeneity among the studies, in addition to heterogeneity measures also, graphical methods are available. Forest plot, L'Abble plot and radial plot are some examples of it. Graphical methods provide us to serve outputs of studies' with their confidence intervals, statistical significance and help us to make decisions about predictions which have been conducted in the studies. [3]

Forest Plot

The first usage of forest plot emerges in 1970s by Freiman et al. He showed the results of studies by using horizontal lines, included with confidence intervals. But it is not a meta-analysis because, for each study, results were served separately. In 1982, Lewis and Ellis found a similar plot to Freiman', but for meta-anaylsis which presents an overall effect on it. [28]

Forest plot is the most commonly used graph type in the meta-analysis. It describes the differences in effect size obtained from the individual studies. Includes effects size predictions, its' 95 percent confidence intervals and odds ratio. The variability among the predictions represents the heterogeneity. The y-axis is reserved for individual studies that are included into the meta-analysis and the x-axis is reserved for logarithmic values of odds ratio or risk ratio. [3, 13]

Following figure 1.1. is an example forest plot. Diamond symbol in the forest plot is the estimated effect size and its' width refers the precision of the estimate. For each study square symbol refers the weight of the study. Bigger squares means higher weight and the line refers the 95% CI for lower/upper limits.

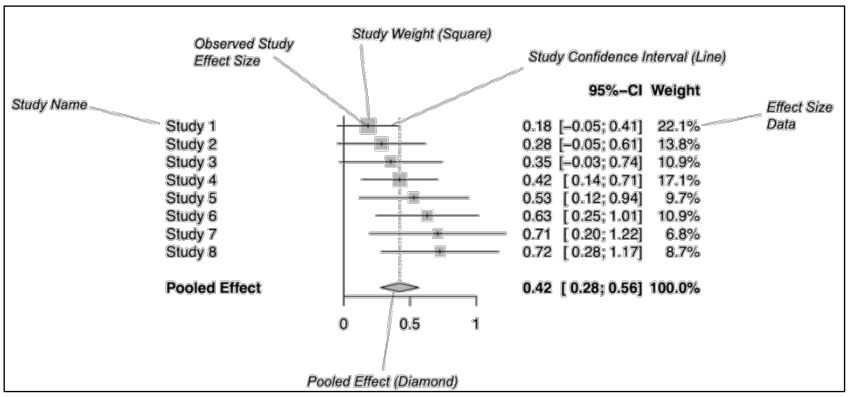

Figure 1.1. Example forest plot [29]

L'Abble Plot

Firstly served in 1987 by Labbe and his friends. L'Abble plot is generated when the result variable is in binary form. Plotted by using the risk measurements of case and control groups. If points are close to the line, included works into the meta-analysis are homogeneous. On the other hand, scattering or large deviations draw attention to heterogeneity. [2,3]

Point size in the graph is related with the weight of that particular study. The points below the line show that the effect size value on the treatment group is higher than the effect size value on the control group, therefore new treatment is less effective than the standard treatment. When the control group and treatment group is compared, the points above the line shows the effectiveness of the new treatment. [2,3]

Radial Plot

The radial plot is a scatter plot, firstly served by Galbraith. Also, known as Galbraith radial graph, Galbraith graph. It is generated by dividing the outputs obtained from each study into the square root of variance ($\frac{\ln(odds\ ratio)}{standart\ error}$) versus the inverse of the standard error. X-axis reserved for inverse of the standard error ($\frac{1}{\sqrt{var(T_i)}}$) and y-axis is reserved for standardized effect size ($\frac{T_i}{\sqrt{var(T_i)}}$), known as z-statistic. In meta-analysis low weighted and high standard error points are located closer to the y-axis. Also, a non-weighted regression line is drawn. If the point is inside the 95 percent confidence interval area, the studies

included into the meta-analysis are homogeneous; outside the 95 percent confidence interval area, the studies included into the meta-analysis are heterogeneous. [3]

COVID-19 is a disease, caused by a virus named SARS-CoV-2 and was firstly reported in December 2019 in Wuhan, China. [30] It has some impacts on many systems in the human body and thyroid is just one of them. Since, it holds a critical role to regulate numerous metabolic processes throughout the body, a few changes in it, may have affect on whole human body. The aim of this meta-analysis was to focus on the effects of COVID-19 on thyroid disorders.

1.11. THYROID

Thyroid is an endocrine gland, situated inferior to the trachea, anterior of the lower neck, below the Adam's apple -called prominentia laryngea in health sciences- and larynx. The shape of it looks like U or H letters and butterfly-shaped. Just weighted 10-15 grams. Contains left and right lobes, each lobe is heightened 4 cm and weighted 2 cm. Isthmus is a narrow tissue that connects the lobes. By swalling it moves up and down. Carotid arters are located on both sides of the thyroid gland. [31] Anatomy of the thyroid gland is shown in the figure 1.2. [31]

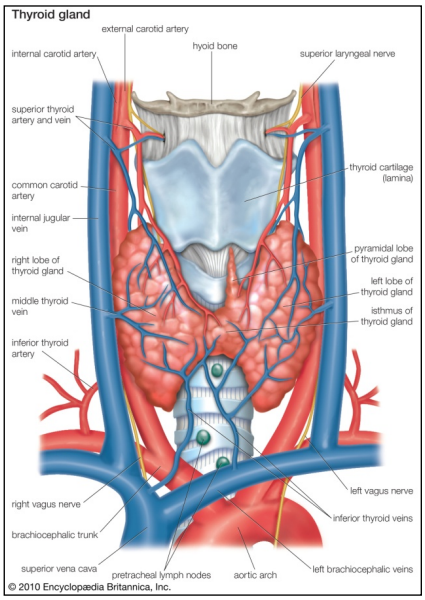

Figure 1.2. Anathomy of thyroid gland [31]

Both lobes contain follicles, those are little globular sacs and follicles contains follicular cells, that are full with a fluid called colloid. This fluid consists of prohormone thyroglobulin. Those cells have some enzymes in order to produce thyroglobulin also, to secrete thyroid hormone from thyroglobulin. [31]

Functionalities and Hormones of the Thyroid Gland

Despite the small size, the functionality of thyroid gland is very vital for the human body and has effects on many activities that occur in the body. Responsible for controlling and releasing hormones which control the metabolism. “When thyroid hormones are needed, thyroglobulin is reabsorbed from the colloid in the follicular lumen into the cells, where it is split into its component parts, including the two thyroid hormones thyroxine (T4) and triiodothyronine (T3)” (“thyroid gland”) [31]. Those hormones are transferred by the blood and stimulate the related organ or system. When they are passed to the blood, 99 percent of them are carried by binding to the carrier proteins which is known as thyroxine-binding globulin (TBG). To observe the effects of the hormones on the tissues, they should be in the free form. But just one percent of the whole hormones in the body are in free form. [32]

Since the thyroid gland works properly, it will secrete sufficient amount of hormone, then metabolism rate will stay fine. Thyroxine is more secreted and produced rather than triiodothyronine by the thyroid gland. Moreover, in many tissues thyroxine hormone can be converted into the triiodothyronine with the support of deiodinases enzyme. Deiodinases enzyme removes one out of four iodine atoms from the thyroxine when it enters the cell. As a result, three iodine atoms remains and triiodothyronine emerges. 80 percent of triiodothyronine is produced from thyroxine by the deiodination process. [31]

Beyond regulating the glucose, lipid and protein metabolism; thyroid hormones are also crucial for working the immune system, regulating the heart rate, regulating the body temperature, brain-intelligence development in fetus and newborn babies, raising and development of children. [33]

Production, secretion and storing functions of T3 and T4 thyroid hormones is controlled by the thyroid-stimulating hormone (TSH). It is secreted from pituitary gland which is located in the brain. TSH is controlled by the thyrotropin-releasing hormone (TRH). It is secreted from the hypothalamus which is located in the brain. Insufficient secretion of thyroid

hormones slows down body functions and surplus secretion of thyroid hormones accelerates the body functions. T3 and T4 are both attached to a protein, but free T3 (FT3) and free T4 (FT4) is not attached to anything. It is the form of hormones which is circulating in the blood.

1.11.1. Thyroid Disorders

Thyrotoxicosis and Hyperthyroidism

Thyrotoxicosis is a case because of the excessive amount of thyroid hormone and any cause may lead to it [34]. Therefore hyperthyroidism is a subset of thyrotoxicosis [35].

Hyperthyroidism is known as overactive thyroid. Excessive production of thyroid hormone has resulted in hyperthyroidism. The prevalence of it is about 2 percent among females and 0.2 percent among males [36]. Too much thyroxine hormone may cause Graves' disease, Plummer's disease and thyroiditis [37]. Increased body metabolism rate and organ activities occur. Its' symptoms are unintentional weight loss, tachycardia, arrhythmia, palpitations, increased appetite, nervousness, anxiety, irritability, tremor, sweating, menstrual changes, sensitivity to heat, diarrhea, goiter, fatigue, tiredness, muscle weakness, sleeping difficulties, skin thinning and brittle-fine hair [37]. There exist some treatment methods to cope with hyperthyroidism. Hormone production can be slowed down by applying anti-thyroid medications. Then, radioactive iodine treatment will apply for definite treatment. In some cases, surgery may be needed to remove the thyroid gland partially or totally [38]. Hyperthyroidism may lead to some complications such as heart problems, brittle bones, eye problems, red-swollen skin and thyrotoxic crisis [37].

Thyroiditis

Thyroiditis is a kind of inflammation occurs in the thyroid gland [39]. Inflammation may resulted as leakage of excess thyroid hormone stored in the thyroid gland into the body [37]. Diagnosed by laboratory blood tests, ultrasonographic-radionuclide imaging and biopsy. Thyroiditis can be examined into two categories: painful and painless. Subacute painful, sporadic painless and acute infectious thyroiditis are types of it. Subacute painful and sporadic painless disorders are known as destruction-induced thyroiditis. The reason

behind the destruction-induced is viral or immune-based attack then, releasing the previously produced hormone into the blood. Destructive attack may lead to pain in subacute thyroiditis but postpartum and sporadic thyroiditis is usually painless. [40]

In 1895 subacute thyroiditis was firstly proposed by Mygind. He reported 18 cases. However, in 1904 the pathology of it was firstly proposed by Fritz de Quervain. Subacute thyroiditis is the most widespread reason for painful thyroid. Also, consists 5 percent of all clinical thyroid disorders. [40]

Specific etiology is unknown but a viral etiology or upper respiratory infection may prodrome to it. Moreover, genetic and autoimmune etiologies are suggested. Triggering the immune system by drugs and vaccines may also have effect on the development of the subacute thyroiditis. [40]

Each year 12.1 cases are diagnosed per 100 thousand people. Incidence is 5 times more in females than males [40]. Usually, pregnant women and children are not affected. Goiter is painful and transient. Inflammation is predominant in left, right or in both lobes. Symptoms are fever, malaise, myalgies and fatique. Initially sudden or gradual pain take place region of thyroid. Pain involves in both lobes and it is constant. In some rare cases, it affects just a particular lobe. Neck pain may confuse with a sore throat. But actually, the pain is in the neck not in the pharynx. It increases by swalling or turning the head. In addition, pain may also be perceived on the chest. [40]

Thyroid antibodies are low titer or absent but it is transient. Thyroglobulin and thyroid peroxidase antibody amounts are generally normal in the blood. Relapse rate is rare. Thyrotoxicosis may be followed by hypothyroidism and the probability of having permanent hypothyroidism is about 5-15 percent. [40] Diagnosed by sedimentation, thyroid function blood tests and radioactive iodine uptake. Thyroid functions turn back to normal in 12-18 months. [40]

Hypothyroidism

Hypothyroidism is known as underactive thyroid. Insufficient production of thyroxine and triiodothyronine hormones are resulted as hypothyroidism. Then, the pituitary gland will increase the TSH secretion. Despite the TSH amount is high, T3 and T4 hormone levels

are low in the blood. Thyroid surgery, radiation therapy, hyperthyroidism treatment, autoimmune diseases, congenital diseases, pituitary disorders, pregnancy, iodine deficiency and some medications may lead to hypothyroidism. [41] Decreased body metabolism rate and organ activities occur. Its' symptoms are fatigue, increased sensitivity to cold, constipation, dry skin, weight gain, puffy face, hoarseness, muscle weakness, elevated blood cholesterol level, muscle aches-tenderness-tiffness, joint pain-stiffness-swelling, heavier than normal or irregular menstrual periods, slowed heart rate, depression, impaired memory, goiter and fine hair. [42] For treatment, different medications are applied according to the cause of hypothyroidism. Hypothyroidism may lead to some complications such as goiter, heart problems, slowed mental functioning, peripheral neuropathy, myxedema, infertility and birth defects. [42]

1.12. COVID-19

COVID-19 is a disease, caused by SARS-CoV-2 virus and was firstly reported in December 2019 in Wuhan, China [30]. World Health Organisation China Country office declared pneumonia cases of unknown cause. In January 2021 noticed that a new coronavirus causes that disease and was early named as 2019-nCoV. After a while called as COVID-19. Despite the other pandemics in recent years around the world; it spread faster and more fatal. Just in three months spread all around the world. On March 12, 2020, epidemic was announced as a pandemic. [43] SARS-CoV-2 is member of coronavirus family which causes respiratory disease and diagnosed by applying a laboratory test. Other instances of coronavirus family are severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS) [44]. Until January 27, 2022 there exist 360.578.392 confirmed cases and 5.620.865 confirmed deaths [45].

SARS-CoV-2 virus enters the body from nose, mouth and eyes. It transmitted by the airbone droplets and infect a human by inhaling the droplets spread by coughing, sneezing, breathing and talking [44]. Also, by touching to contaminated surfaces by the virus then, taking the hands to nose, mouth and eyes [46]. When virus enters the body, it passes through the nasal channels and mucous membrane located in the back of the throat. Later, it binds the cells and begins to multiply itself. Firstly, travels through the lungs tissue then, spread to other tissues as well [44].

The most common symptoms are cough, fever and shortness of breath. Other common symptoms are loss of taste or smell, tiredness, body/muscle aches, congestion/runny nose, sore throat, headaches, diarrhea, vomiting/nausea. Children and adults have similar symptoms but usually milder in children. After virus enters the body, symptoms may appear within two to fourteen days [44]. In order to relieve/reduce the symptoms take pain killers, have rest and increase fluid intake.

Treatment of COVID-19 may be provided by several methods such as using drugs, monoclonal antibody medication and plasma therapy. U.S. Food and Drug Administration (FDA) approved the usage of the drugs remdesivir (Veklury), Paxlovid, molnupiravir and baricitinib (Olumiant) to treat COVID-19 cases [47]

The higher risk group affected by the COVID-19 are ageing over sixty, having serious chronic medical conditions, heart disease, hypertension, diabetes, chronic respiratory disease, cancer, being a healthcare employee etc [46].

1.13. THYROID AND COVID-19 RELATION

Thyroid hormones have a critical role in innate immune response regulation. Excessive or fewer amounts of thyroid hormones may lead to innate immune response dysregulation. According to Hariyanto's (2020) meta-analysis study, innate immune responses were thought to be contributed the most to the pathogenesis of COVID-19 as it is the front lines of the body defence system to fight against SARS-CoV-2 virus. Any dysregulation in innate immune response exhibited by increase in CD14+ monocytes, macrophages, neutrophil and decrease in natural killer (NK) cells level and increased levels of complement was associated with more severe COVID-19 disease. Moreover, an increase in levels of IL-6 and TNF-α –those are the proinflammatory cytokines- was observed in thyroid-disordered individuals. Those cytokines also exist in severe COVID-19 patients. Treatment of some thyroid diseases, corticosteroids are used as medications. But corticosteroid treatment was associated with longer hospitalization stays and higher mortality rates caused by COVID-19. [48]

Accroding to Trimboli (2021), existing of receptor for SARS-CoV-2 entry in thyroid cells has been proven. Also, "have several clues that SARS-CoV-2 can have effects on the

thyroid gland. High levels of expression of ACE2 receptor and transmembrane protease serine 2 have been found in thyroid cells, abnormal immune responses and cytokine storm associated to COVID-19 may induce thyroid gland inflammation.” [49]

2. METHODOLOGY

As mentioned in the introduction chapter there exist 8 steps in order to conduct a meta-analysis. The methodology chapter will serve the first 4 steps. These are identifying the problem; determining the inclusion and exclusion criteria; completing the literature review according to the inclusion criteria; extracting, clustering and coding the studies in terms of methodology and outputs steps.

2.1. IDENTIFYING THE PROBLEM

For treatment and prevention from COVID-19, lots of studies have been done. Also, to observe the relationship between COVID-19 and some other diseases. During the pandemic, thyroid disorders get increased, and those disorders are usually interpreted as autoimmune mechanisms.

Thyroid disorders are conditions that affect the thyroid gland. The thyroid gland has a critical role to regulate numerous metabolic processes throughout the body. Different types of thyroid disorders affect either its' structure or function. The thyroid gland produces hormones that help control many vital functions of the body. When it is not working properly, it may have an impact on the entire body. If the body makes too much or too little thyroid hormone, there can be serious conditions that need to be treated.

Due to the limitations in terms of time, some studies have been done with insufficient or unbalanced sample groups, populations, demographic structures etc. Resulted as lower reliability level. Meta-analysis is one of the best techniques to come up with higher evidence levels. It combines the results of multiple scientific studies and generates outputs from previous studies by using statistical analysis.

2.2. INCLUSION AND EXCLUSION CRITERIA

As it is served in the previous section there exist some advantages to sum up similar outcomed multiple studies which have been done on the same topic. For instance, increased reliability level and increased sample size to classify the studies either

statistically significant or not. In order to extract similar outcomed studies after having the literature review, there should be some criteria. Those criterias are aimed to minimize the differences among the studies and provide combining the findings effortlessly. [50] The criterias which have to be in a study -added into the meta-analysis- defined as inclusion criteria. For this work inclusion criterias identified as followings:

- Published in the English language.
- Published after January 1, 2019. Since, the COVID-19 pandemic spread out in 2019.
- Sample groups should not have known thyroid dysfunction before.
- Blood test results for thyroid function tests (TFT) with median, mean, standard deviation or interquartile range (IQR) values.
- More than one severity classification for the COVID-19 patients.
- The category of the study should be retrospective cohort, prospective cohort, cross-sectional or clinical trial.

As well as the inclusion criteria there should be some exclusion criterias if the studies do not meet the aim of the meta-analysis subject or do not contain statistical data etc. The reason of exclusion for each reviewed study is explained in the literature review section. For this work exclusion criterias identified as followings:

- The age of the sample groups should be greater than 18.
- Pregnant sample groups.
- Only male or female sample groups.
- Review, case report, meta-analysis, editorial, follow-up studies.
- No existence of full-text study.

2.3. LITERATURE REVIEW

After determining the inclusion and exclusion criterias the next step for a meta-analysis is the literature review. Since it is based on the outputs of previous works, a systematic search should be conducted. A literature review was completed on PubMed [51] and Science Direct [52] databases. Those are the free search engines which are especially for science and biomedical studies.

Aiming to discover similar resulted studies, a literature review should be done after identifying some keywords. For this study key words are "hypothyroidism", "hyperthyroidism”, "thyroiditis", "subacute thyroiditis”, "autoimmune thyroiditis", "hyperthyroxinemia", "T3", "T4", "TSH", "COVID-19", "SARS-CoV-2", "coronavirus disease 2019", "mortality", "ICU admission". Both PubMed and Science Direct search engines are able to accept boolean operators (or, and, not) between the keywords. By using the keywords and boolean expressions one or more search queries should be prepared, then executed. PubMed accepts unlimited number of boolean operators, despite that the maximum number of boolean operators in Science Direct is 9.

Regarding to keywords, the search query for the literature review in PubMed is determined as (("hypothyroidism"[Mesh] OR "hyperthyroidism"[Mesh] OR "thyroiditis" OR "subacute thyroiditis" OR "autoimmune thyroiditis" OR "hyperthyroxinemia" OR "T3" OR "T4" OR "TSH") AND ("COVID-19"[Mesh] OR "SARS-CoV-2" OR "COVID-19" OR "coronavirus disease 2019" OR "mortality" OR "ICU admission")). MeSH (Medical Subject Headings) is a term using for indexing the articles in PubMed.

The query was executed on April 9, 2022. Search is filtered as published in 2019 or after, English language restricted, abstract, and free full-text availability. After applying those filters 1174 studies exists in the search result. Then the result was exported in CSV format into an Excel sheet with PMID, title, author, citation, first author name, journal/book, publication year/date, PMCID and DOI number.

3-phased review procedure was applied on the studies. All studies were reviewed by two examiners. The first phase was reviewed regarding to the title and study design (meta-analysis, review, editorial, case report etc.) category declared in the PubMed. Table 2.1 presents the number of studies for each category. 1126 studies were excluded because of the unrelated subject or inappropriate study design.

60 of them are related with COVID-19 vaccine, 20 of them are meta-analysis, 267 of them are related with cancer, 502 of them do not contain any word ending with “19” or starting with “sars”, 190 of them do not contain any word started with “thyro” or “endo”, 53 of them is review, 2 of them is editorial, 31 of them is case report and 1 of them is follow-up studies.

Table 2.1. Categorization of studies from PubMed according to title of the studies

Category	Frequency
Related with COVID-19 vaccine	60
Meta-analysis study	20
Related with cancer	267
not contain *19 or SARS* words	502
not contain thyro* or endo* words	190
Review study	53
Editorial study	2
Case report study	31
Follow-up study	1

In the second phase, the abstracts of the remaining 48 studies are read. Table 2.2 presents the number of studies for each category. 32 study was excluded because of inappropriate sample group/study design, already having thyroid disorders or do not stating the exclusion of thyroid disordered patients. 2 of them is not classified as case report in the PubMed but the sample size is less than 5 patients, so marked as a case report. Similar to that case 1 of them is review and 2 of them is follow-up study. Seventeen of them having thyroid dysfunction is not declared as exclusion criteria, 5 of them were conducted on child, pregnant or only male/female sample groups and 5 of them does not give two or more COVID-19 severity.

Table 2.2. Categorization of studies from PubMed according to abstract of the studies

Category	Frequency
Case report	2
Review study	1
Follow-up study	2
Not exclusion of known thyroid disorders	17
Inappropriate sample group	5
Severity of COVID-19 not given	5

In the third phase, full-text of the remaining 16 studies are read. 4 of them were excluded since the severity of COVID-19 is not given. After 3-phased review, the remaining 12 studies included into the meta-analysis.

Table 2.3. Categorization of studies from PubMed according to full-text of the studies

Category	Frequency
Severity of COVID-19 not given	4

Regarding to keywords, search query for the literature review in Science Direct is divided into 3 subqueries. Because it is able to accept maximum 9 boolean operators. Queries are prepared with the same meaning logic with the PubMed. First query is (("hypothyroidism"[Mesh] OR "hyperthyroidism"[Mesh] OR "thyroiditis") AND ("COVID-19"[Mesh] OR "SARS-CoV-2" OR "COVID-19" OR "coronavirus disease 2019" OR "mortality" OR "icu admission")). Second query is (("subacute thyroiditis" OR "autoimmune thyroiditis" OR "hyperthyroxinemia) AND ("COVID-19"[Mesh] OR "SARS-CoV-2" OR "COVID-19" OR "coronavirus disease 2019" OR "mortality" OR "ICU admission")). Third query is (("T3" OR "T4" OR "TSH") AND ("COVID-19"[Mesh] OR "SARS-CoV-2" OR "COVID-19" OR "coronavirus disease 2019" OR "mortality" OR "ICU admission"))

The query was executed on May 3, 2022. Search is filtered as published in 2019 or after. English language restriction and full-text availability can not be applied as filter before the search, so after exporting the results they are excluded. Consecutively 1036, 122, 1928 (in total 3086) studies take place as the results in queries. The results were exported to "Mendeley Reference Manager", then converted to xml format and get into an Excel sheet with source type, title, year, standard number, pages, journal name, volume, issue, first author name and URL.

Same as PubMed 3-phased review procedure was applied on to the studies. All studies were reviewed by two examiners. In the first phase reviewed regarding to duplicates from PubMed, duplicate with any of the Science Direct queries, title, study design and language. Table 2.4 presents the number of studies for each category. 3064 studies were excluded because of duplicate, inappropriate study designs.

354 of them is duplicate with any other query in Science Direct, 50 of them is duplicate with PubMed query, 210 of them is related with COVID-19 vaccine, 20 of them is meta-analysis, 135 of them is related with cancer, 1098 of them does not contain any word ending with "19" or starting with "sars", 1025 of them does not contain any word started

with “thyro” or “endo”, 150 of them is review, 8 of them is case report, 8 of them is conference abstract and 6 of them is not published in English.

Table 2.4. Categorization of studies from Science Direct according to title of the studies

Category	Frequency
Duplicate with other Science Direct	354
Duplicate with PubMed	50
Related with COVID-19 vaccine	210
Meta-analysis study	20
Related with cancer	135
not contain *19 or SARS* words	1098
not contain thyro* or endo* words	1025
Review study	150
Case report study	8
Conference abstract	8
Not English	6

In the second phase, the abstracts of the remaining 22 studies are read. Table 2.5 presents the number of studies for each category. Eighteen study was excluded because of inappropriate study design and unrelated subject. Three of them is not classified as case report in Science Direct but the sample size is 1 patient, so marked as a case report. Four of them is review study but it is not stated in the title so, they are screened. Similar to that case 1 of them is meta-analysis. Ten of them are about other subjects about thyroid disorders.

Table 2.5. Categorization of studies from Science Direct according to abstract of the studies

Category	Frequency
Case report study	3
Review study	4
Meta-analysis study	1
Unrelated subject	10

In the third phase, full-text of the remaining 4 studies are read. 2 of them were excluded since thyroid function test (TFT) is not given and 1 of them is a follow-up study. After 3-phased reviewed, the remaining one study included into the meta-analysis.

Table 2.6. Categorization of studies from Science Direct according to full-text of the studies

Category	Frequency
Severity of COVID-19 not given	2
Follow-up study	1

All in all, 12 studies from PubMed and 1 study from Science Direct –total 13 studies- search engines are meet the inclusion criteria then, included into the meta-analysis. Data collection procedures explained until here, are summarised in the following flowchart (Figure 2.1). Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) ckecklist steps are followed.

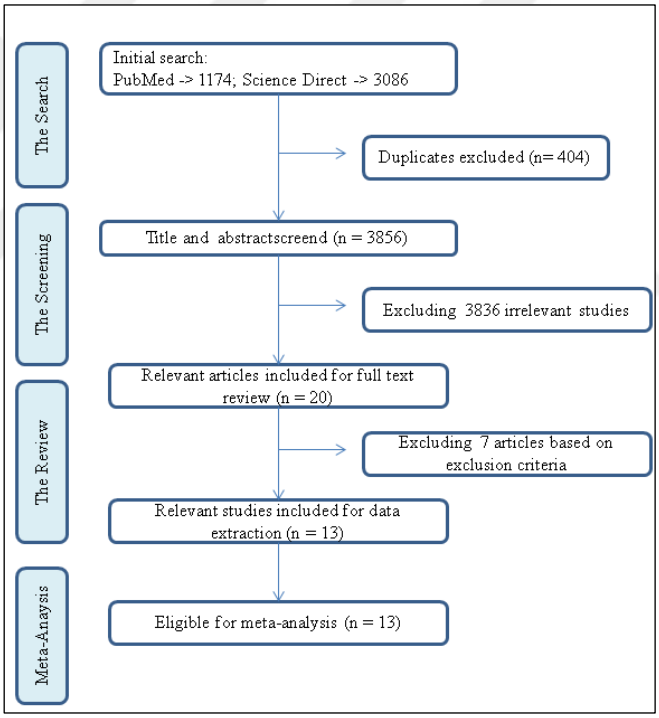

Figure 2.1. PRISMA flow-chart

Table 2.7 shows the included studies into the meta-analysis with the first author, article name, year, and database information.

Table 2.7. Summary of included studies into the meta-analysis

Article Number	Article Name	First Author	Publication Year	Database
1	Thyroid Hormone Profile and Its Prognostic Impact on the Coronavirus Disease 2019 in Korean Patients [53]	Seo Young Sohn	2021	PubMed
2	Thyroid hormone concentrations in severely or critically ill patients with COVID-19 [54]	Weibo Gao	2020	PubMed
3	Thyroid hormone alterations in critically and non-critically ill patients with SARS-CoV-2 infection [55]	Dimitra Argyro Vassiliadi	2021	PubMed
5	Thyroid function analysis in COVID-19: A retrospective study from a single center [56]	Jahanzeb Malik	2021	PubMed
6	Thyroid Dysfunction in Relation to Immune Profile, Disease Status, and Outcome in 191 Patients with COVID-19 [57]	David Tak Wai Lui	2020	PubMed
7	Thyroid dysfunction in COVID-19 patients [58]	R. Baldelli	2021	PubMed
8	The Prognostic Role of Metabolic and Endocrine Parameters for the Clinical Severity of COVID-19 [59]	Shokoufeh Bonakdaran	2022	PubMed
9	The prognostic impact of thyroid disorders on the clinical severity of COVID-19: Results of single-centre pandemic hospital [60]	Mehmet Güven	2021	PubMed
12	The Association of Thyroid Hormone Changes with Inflammatory Status and Prognosis in COVID-19 [61]	Ceyda Dincer Yazan	2021	PubMed
13	Role of non-thyroidal illness syndrome in predicting adverse outcomes in COVID-19 patients predominantly of mild-tomoderate severity [62]	David Tak Wai Lui	2021	PubMed
14	Thyroid Function Test in COVID-19 Patients: A Cross-Sectional Study in a Tertiary Care Hospital [63]	Kaushik Sen	2021	PubMed
15	Analysis of thyroid function in Japanese patients with coronavirus disease 2019 [64]	Sachiko Nakamura	2021	PubMed
20	Thyroid Function Abnormalities in COVID-19 Patients [65]	Weibin Wang	2021	Science Direct

2.3.1. Extracting, Clustering and Coding the Studies

The fourth step in order to conduct a meta-analysis is extracting, clustering and coding the studies which are assessed from the literature review. Table 2.8 presents the study design, country, intervention group sample size, control group sample size, age with mean ± standard deviation and male ratio among the 13 selected studies. In summary, the most used study design is cohort (76.92%), most of the observations take place in China (30.77%) with 1640 patients, average sample size is 217.6 patients, average age is 58.73 with the standard deviation 15.63 and 58.79% of the patients are male. In total, 2829 individuals were conducted into the meta-analysis.

Table 2.8. Selected articles overall details

Article Number	Study Design	Country	Intervention Group Sample Size	Control Group Sample Size	Age (mean ± SD)	Male (%)
1	Retrospective cohort	South Korea	119	N/A	64.3 ± 16.8	52.1
2	Retrospective cohort	China	100	N/A	62.59 ± 14.01	50
3	Prospective cohort	Greece	102	94	59.3 ± 18.3	66.3
5	Retrospective cohort	Pakistan	48	28	56.1 ± 16.4	60.53
6	Prospective cohort	China	191	N/A	53.5 ± 17.2	51.8
7	Retrospective cohort	Italy	46	20	59.6 ± 14.55	69.56
8	Cross-sectional	Iran	70	N/A	61.0 ± 15.3	55.7
9	Prospective cohort	Turkey	250	N/A	66.6 ± 17.9	63
12	N/A	Turkey	205	N/A	57.45 ± 13.29	55.12
13	Prospective cohort	China	367	N/A	52.25 ± 20.09	46.9
14	Cross-sectional	India	60	N/A	N/A	N/A
15	Retrospective cohort	Japan	147	N/A	67.8 ± 13.07	64.6
20	Retrospective cohort	China	84	898	57.89 ± 13.47	63

Regarding to the published year, 76.92% of the selected studies are published in 2021, 15.38% in 2020 and 7.69% in 2022. Table 2.9 serves the frequencies for each study' published year.

Table 2.9. Published year-based selected article frequencies

Published Year	Frequency of the Studies	Percentage (%)
2020	2	15.38
2021	10	76.92
2022	1	7.69

Regarding to the study type 76.92% of the selected studies are built by cohort studies. Cohort studies are divided into 2 sub-groups: retrospective cohort and prospective cohort. Retrospective cohort study ratio is 60% and prospective cohort study ratio is 40%. 15.38% of the studies are cross-sectional and for 1 study (7.69%) study type is not declared. Table 2.10 serves the frequencies for each study design.

Table 2.10. Study design-based selected article frequencies

Study Design	Frequency of the Studies	Percentage (%)
Cohort	10	76.92
Retrospective Cohort	6	60
Prospective Cohort	4	40
Cross-Sectional	2	15.38
Not Given	1	7.69

Regarding to the country of the studies, most of the studies were completed in China with ratio 30.77%, then in Turkey with 15.38%, other countries are Greece, India, Iran, Italy, Japan, Pakistan, South Korea conducted 1 study (7.69%) per each. Table 2.11 serves the frequencies for each study country.

Table 2.11. Country-based selected article frequencies

Country	Frequency of the Studies	Percentage (%)
China	4	30.77
Turkey	2	15.38
Greece	1	7.69
India	1	7.69
Iran	1	7.69
Italy	1	7.69
Japan	1	7.69
Pakistan	1	7.69
South Korea	1	7.69

Intervention group is assumed as COVID-19 patients. Control group is assumed as both non-COVID-19 patients and healthy individuals. Sample size of intervention group is 1789 individuals, with average 137.62 and 93.83 standard deviation. Sample size of control group is 1040 individuals, with average 260 and 426.6 standard deviation. In total, 2829 individuals were conducted into the meta-analysis. The most population included into the meta-analysis is from China (57.97%), Turkey (16.08%) and Greece (6.93%). Table 2.12 serves the country-based intervention and control populations.

Table 2.12. Country-based intervention and control groups sample sizes

Country	Intervention Group Sample Size (n, %)	Control Group Sample Size (n, %)	Total Sample Size (n, %)
China	742 (41.48)	898 (86.35)	1640 (57.97)
Turkey	455 (25.43)	N/A	455 (16.08)
Greece	102 (5.70)	94 (9.04)	196 (6.93)
India	60 (3.35)	N/A	60 (2.12)
Iran	70 (3.91)	N/A	70 (2.47)
Italy	46 (2.57)	20 (1.92)	66 (2.33)
Japan	147 (8.22)	N/A	147 (5.20)
Pakistan	48 (2.68)	28 (2.69)	76 (2.69)
South Korea	119 (6.65)	N/A	119 (4.21)

Age and gender variables are calculated in terms of weighted mean for both intervention and sample groups. Regards to the total sample size weighted mean age is 58.73 and standard deviation is 15.63. Weighted mean ratio for male sample size is 58.79% and female sample size is 41.21%.

As it is defined in the inclusion criteria all selected articles should contain TFT and at least two severity groups. In addition to this, 3 articles serve comorbidity, 11 articles serve thyroid dysfunction, 8 articles serves ICU admission, 4 articles contain control group with non-COVID-19 patients and healthy individuals. In order to summarize the severity outcome, 2 groups were determined as non-severe and severe-to-critical. Mild, moderate, non-ICU admission patients were added into non-severe group and severe, critical, ICU admission, ward patients added into severe-to-critical group. Table 2.13 serves the TFT, comorbidity, thyroid dysfunction, mortality, ICU admission, control group existence status (+: exist; -; not exist) and severity group for each selected study.

Table 2.13. Article-based given properties

Article Number	TFT	Comorbidity	Thyroid Dysfunction	Mortality	ICU Admission	Control Group	Severity
1	+	-	+	+	+	-	non-severe, severe-to-critical
2	+	-	-	+	-	-	non-severe, severe-to-critical
3	+	-	+	-	+	+	ICU, Ward, outpatients
5	+	-	+	-	+	+	moderate, severe, critical
6	+	-	+	-	-	-	mild, moderate, severe
7	+	-	+	-	+	+	hospitalized, ICU
8	+	+	-	+	+	-	ICU admission and non-ICU admission
9	+	-	+	+	+	-	ICU admission and non-ICU admission
12	+	-	+	+	+	-	ICU admission and non-ICU admission
13	+	+	+	-	-	-	Mild, moderate, severe
14	+	-	+	-	-	-	Mild, moderate, severe
15	+	+	+	-	+	-	Mild, moderate, severe(ICU admission)
20	+	-	+	-	-	+	mild/moderate, severe/critical

TFT exists as TSH, FT3, TT3, FT4, TT4 values in the selected articles. Nearly all articles serve only some of the TFT's. All of them present TSH, 8 of them FT3, 6 of them TT3, 11 of them FT4 and 3 of them TT3 values. Each TFT is given in different units, then all values converrrttted into the most used unit. For TSH μIU/mL, for FT3 pg/ml, for TT3 ng/dL, for FT4 ng/dL, for TT4 mcg/dL units are used. Table 2.14 presents the article-based TFT and its' units. "X" signs correspond the existence.

Table 2.14. Article-based TFT

Article Number	TSH	FT3			TT3			FT4		TT4	
	µIU/mL	pmol/L	**pg/ml**	**ng/L**	nmol/L	**ng/dL**	ng/ml	pmol/L	**ng/dL**	nmol/L	**mcg/dL**
1	X						X		X		
2	X	X						X			
3	X					X			X		
5	X					X					X
6	X	X						X			
7	X		X						X		
8	X					X			X		
9	X		X						X		
12	X			X					X		
13	X	X						X			
14	X		X				X		X		X
15	X		X						X		
20	X				X					X	

Mean and standard deviation (SD); median and interquartile range (IQR); or median and minimum-maximum values are served in the articles to demonstrate TFT's. All values should be in the mean ± SD or number (n) form to include into the meta-analysis. Firstly, all median (IQR) and median (min-max) values are translated into mean and SD by using an online estimator [66]. After that, an online weighted average calculator [67] was used for estimating weighted average. Weighted average is a method that emerges from the multiplication of each value by its frequency. Despite this, simple mean assumes that each value is assigned to equal weights. Instead of simple mean, using the weighted mean is more accurate, since each study has different sample sizes.

Outputs of each study were extracted in terms of numerical values (mean ± SD or number (n)). Similar output works were grouped aiming to include into the meta-analysis. Since all studies resulted in different outputs; TFT vs COVID-19 severity, comorbidity vs COVID-19 severity, TFT vs mortality, thyroid dysfunction vs COVID-19 severity and TFT vs COVID-19 existence meta-analysis topics are determined.

TFT vs COVID-19 severity subject consists of 5 meta-analyses inside it and aims to observe the TSH, FT3, TT3, FT4, TT4 values for mild/moderate and severe/critical COVID-19 severity patients. Consecutively 9, 5, 5, 7, 2 articles were included.

Comorbidity vs COVID-19 severity subject consist of 2 meta-analyses inside it and aims to observe diabetes and hypertension diseases for mild/moderate and severe/critical COVID-19 severity patients. 2 articles are included per each disease.

TFT vs mortality subject consists of 4 meta-analyses inside it and aims to observe the deceased/survived COVID-19 patient amounts according to the TSH, FT3, TT3, FT4 TFT values. Consecutively 4, 3, 3, 5 articles were included.

Thyroid dysfunction vs COVID-19 severity subject consists of 1 meta-analysis inside it and aims to observe the thyroid dysfunction for mild/moderate and severe/critical COVID-19 severity patients. Thyroid dysfunction evaluated as non-thyroidal illness syndrome (NTIS), subclinical thyrotoxicosis, thyrotoxicosis, overt thyrotoxicosis, hypothyroidism, overt hypothyroidism, subclinical hypothyroidism, subnormal TSH, low FT3, high FT3, low FT4 and high FT4. 7 articles were included.

TFT vs COVID-19 existence subject consists of 4 meta-analysis inside it and aims to observe the TSH, TT3, TT4 values for COVID-19 and non-COVID-19 patients. In addition, TSH value was investigated on COVID-19 and healthy individuals. Consecutively 3, 3, 2, 2 articles were included.

2.4. SOFTWARE USED FOR ANALYSING THE DATA

There are different software and tools to complete a meta-analysis. In this study, GRADEpro [68], Cochrane's Review Manager (RevMan) version 5.4.1, MedCalc Statistical Software version 18.6. GRADEpro Guideline Development Tool (GDT) is an online software aiming to create Summary of Findings (SoF) tables in Cochrane systematic reviews [69]. Quality score of the studies were calculated by using it. RevMan was preferred because of ease of use and free to use. Heterogeneity tests and forest plot were conducted in RevMan. Different from RevMan, MedCalc is licensed software and has to pay a licence fee. Publication bias among the studies conducted in the meta-analysis was observed by MedCalc. Funnel plot, Egger's and Begg's tests were generated in it.

2.5. QUALITY SCORE ASSESSMENT OF THE STUDIES

GRADEpro online tool was used to make the quality score assessment of studies. Quality score of each meta-analysis was calculated by taking into consideration the study design risk of bias, inconsistency, indirectness, imprecision, number of patients and effect size. Quality score of all meta-analyses was resulted as “low”. Including observational studies (cohort, cross-sectional) and due to the limitations in terms of time, some studies have been done with insufficient or unbalanced sample groups, populations, demographic structures etc. resulted as lower reliability levels. Appendix A serves the quality assessment for each meta-analysis.

3. ANALYSIS AND RESULTS

As mentioned in the introduction chapter there exists 8 steps to conduct a meta-analysis. In the previous section methodology chapter served the first 4 steps. This chapter will introduce the last 4 steps. Which are applying statistical techniques; combining the outputs of the studies; establishing the relationship between the outputs of the studies and characteristics of meta-analysis; re-reviewing, interpreting, and presenting the results

Aiming to explore the effects of COVID-19 on thyroid disorder, 5 main meta-analysis topics are identified. Due to each study conducted into the meta-analysis containing different types of numerical outputs, some sub-meta-analysis was applied. In total 16 meta-analysis were implemented.

TFT and COVID-19 Severity

- Meta-Analysis 1: TSH (μIU/mL) vs COVID-19 Severity
- Meta-Analysis 2: FT3 (pg/ml) vs COVID-19 Severity
- Meta-Analysis 3: TT3 (ng/dL) vs COVID-19 Severity
- Meta-Analysis 4: FT4 (ng/dL) vs COVID-19 Severity
- Meta-Analysis 5: TT4 (mcg/dL) vs COVID-19 Severity

Comorbidity and COVID-19 Severity

- Meta-Analysis 6: Comorbidity (Diabetes) vs COVID-19 Severity
- Meta-Analysis 7: Comorbidity (Hypertension) vs COVID-19 Severity

TFT and Mortality Rates by COVID-19

- Meta-Analysis 8: TSH (μIU/mL) vs Mortality
- Meta-Analysis 9: FT3 (pg/mL) vs Mortality
- Meta-Analysis 10: TT3 (ng/dL) vs Mortality
- Meta-Analysis 11: FT4 (ng/dL) vs Mortality

Thyroid Dysfunction and COVID-19 Severity

- Meta-Analysis 12: Thyroid Dysfunction vs COVID-19 Severity

TFT and COVID-19 Existence

- Meta-Analysis 13: TSH (μIU/mL) vs COVID-19 Existence
- Meta-Analysis 14: TT3 (ng/dL) vs COVID-19 Existence
- Meta-Analysis 15: TT4 (mcg/dL) vs COVID-19 Existence
- Meta-Analysis 16: TSH (μIU/mL) vs COVID-19/Healthy Status

3.1. SUMMARIES OF THE STUDIES INCLUDED IN THE META-ANALYSIS

3.1.1. Study by Sohn (2021): Thyroid Hormone Profile and Its Prognostic Impact on the Coronavirus Disease 2019 in Korean Patients [53]

This retrospective cohort study was conducted in South Korea on 119 patients with COVID-19. Disease severity (non-severe and severe to critical) and mortality based on TFT (TSH, TT3, FT4) were given, expressed as median (IQR). Also, thyroid dysfunctioned (NTIS, subclinical thyrotoxicosis, subclinical hypothyroidism) number of patients are given based on disease severity and mortality, expressed as number (%). TSH and TT3 levels were lower in patients at severe to critical severity group than in non-severe group ($p < 0.05$). In conclusion, thyroid dysfunction is common in COVID-19 patients. TSH and TT3 levels may be important indicators of COVID-19 severity. Low TT3 levels may have a prognostic significant in COVID-19 related outcomes.

3.1.2. Study by Gao (2020): Thyroid hormone concentrations in severely or critically ill patients with COVID-19 [54]

This retrospective cohort study was conducted in China on 100 patients with COVID-19. Disease severity (non-severe and severe to critical) and mortality based on TFT (TSH, FT3, FT4) were given, expressed as mean (SD). FT3 levels are significantly lower in severe COVID-19 patients than in non-severe patients.

3.1.3. Study by Vassiliadi (2021): Thyroid hormone alterations in critically and non-critically ill patients with SARS-CoV-2 infection [55]

This prospective cohort study was conducted in Greece on 196 patients. 102 of them are COVID-19 and 94 of them are non-COVID-19 patients. Disease severity (ICU, ward, outpatients) based on TFT (TSH, TT3, FT4) was given, expressed as mean (SD). Also, thyroid dysfunctioned (NTIS, thyrotoxicosis, overt thyrotoxicosis, subclinical thyrotoxicosis, hypothyroidism, overt hypothyroidism, subclinical hypothyroidism) number of patients are given based on disease severity and mortality, expressed as number (%). TT3 levels are lower in ICU patients rather than in ward and outpatients. Low FT4 and low TSH levels are more common in ICU patients. The frequency of NTIS is independent from COVID-19 but related with disease severity.

3.1.4. Study by Malik (2021): Thyroid function analysis in COVID-19: A retrospective study from a single center [56]

This retrospective cohort study was conducted in Pakistan on 76 patients. 48 of them are COVID-19 and 28 of them are non-COVID-19 patients. Non-COVID-19 patients have pneumonia. Disease severity (moderate, severe, and critical) based on TFT (TSH, TT3, TT4) are given, expressed as mean (SD). TSH levels are higher in moderate and critical COVID-19 patients rather than non-COVID-19 patients. TT3 levels are lower in COVID-19 pneumonia patients than in non-COVID-19 pneumonia patients.

3.1.5. Study by Lui (2020): Thyroid Dysfunction in Relation to Immune Profile, Disease Status, and Outcome in 191 Patients with COVID-19 [57]

This prospective cohort study was conducted in China on 191 patients with COVID-19. Thyroid dysfunctioned (normal TSH, subnormal TSH, normal FT3, low FT3) number of patients is given based on disease severity (mild, moderate, severe), expressed as number (%). Thirteen percent of patients having thyroid dysfunction are commonly in mild to moderate COVID-19 severity.

3.1.6. Study by Baldelli (2021): Thyroid dysfunction in COVID-19 patients [58]

This retrospective cohort study was conducted in Italy on 66 patients. 46 of them were COVID-19 and 20 of them were healthy. Disease severity (hospitalized and ICU) based on TFT (TSH, FT3, FT4) was given, expressed as mean (SD). FT3 and TSH levels in COVID-19 patients are lower than in healthy individuals. Among the COVID-19 patients, ICU patients have lower FT3 and TSH levels.

3.1.7. Study by Bonakdaran (2022): The Prognostic Role of Metabolic and Endocrine Parameters for the Clinical Severity of COVID-19 [59]

This cross-sectional study was conducted in Iran on 70 patients with COVID-19. Disease severity (ICU and non-ICU) and mortality based on TFT (TSH, TT3, FT4) are given, expressed as mean (SD). Also, disease severity-based comorbidity (diabetes, CVD, hypertension) amounts are given, expressed as number (%). Mortality rates are higher in with patients low TT3 levels.

3.1.8. Study by Güven (2021): The prognostic impact of thyroid disorders on the clinical severity of COVID-19: Results of single-centre pandemic hospital [60]

This prospective study was conducted in Turkey on 250 patients with COVID-19. Disease severity (ICU and non-ICU) and mortality based on TFT (TSH, FT3, FT4) are given, expressed as median (IQR). Also, number of thyroid dysfunctioned (thyroid nodule, overt thyroid disorder, total hypothyroidism, overt hypothyroidism, subclinical hypothyroidism, total thyrotoxicosis, overt thyrotoxicosis, subclinical thyrotoxicosis, NTIS) patients were given based on disease severity, and mortality. Overt thyroid disorders are more common in critical severity patients. FT4 levels of ICU patients are lower than non-ICU admitted patients. No significant difference in TSH levels among the survived and deceased patients but FT4 and FT3 levels of deceased patients are significantly lower than survived individuals.

3.1.9. Study by Yazan (2021): The Association of Thyroid Hormone Changes with Inflammatory Status and Prognosis in COVID-19 [61]

This study was conducted in Turkey on 205 patients with COVID-19. Disease severity (ICU and non-ICU) and mortality based on TFT (TSH, FT3, FT4) are given, expressed as median (min-max). ICU admission rates are lower in euthyroid patients and none of the euthyroid patients died. Patients with lower FT3, higher FT4, lower TSH have higher mortality rates.

3.1.10. Study by Lui (2021): Role of non-thyroidal illness syndrome in predicting adverse outcomes in COVID-19 patients predominantly of mild-to-moderate severity [62]

This prospective cohort study was conducted in China on 367 patients with COVID-19. Thyroid dysfunctioned (normal TFT and NTIS) number of patients were given based on disease severity (mild, moderate, severe), expressed as number (%). NTIS is not common in mild-to-moderate severity patients.

3.1.11. Study by Sen (2021): Thyroid Function Test in COVID-19 Patients: A Cross-Sectional Study in a Tertiary Care Hospital [63]

This cross-sectional study was conducted in India on 60 patients with COVID-19. Disease severity (mild, moderate, severe) based on TFT (TSH, FT3, FT4, TT3, TT4) is given, expressed as median (IQR). There is no significant difference in TFT values regards to disease severity.

3.1.12. Study by Nakamura (2021): Analysis of thyroid function in Japanese patients with coronavirus disease 2019 [64]

This retrospective cohort study was conducted in Japan on 147 patients with COVID-19. Disease severity (mild, moderate, severe) based on TFT (TSH, FT3, FT4) is given, expressed as median (IQR). Disease severity-based comorbidity (diabetes, hypertension,

heart failure, chronic kidney disease, malignant tumour) amounts are given, expressed as number (%). Also, thyroid dysfunctioned (low FT3, high FT3) number of patients are given based on disease severity, expressed as number (%). More than 70% of patients having moderate to severe disease have low FT3 levels.TSH and FT3 levels are inversely proportional to the disease severity. Low FT3 level patients have higher mortality rates rather than normal FT3 level patients.

3.1.13. Study by Wang (2021): Thyroid Function Abnormalities in COVID-19 Patients [65]

This retrospective cohort study was conducted in China on 982 patients. 84 of them are COVID-19, 91 of them are non-COVID-19 but having pneumonia and 807 of them are healthy. TFT values (TSH, TT3, TT4) are given for those 3 sample groups. Also, thyroid dysfunctioned number of patients is given based on disease severity (mild/moderate and severe/critical), expressed as number (%). In severe patients, thyroid dysfunction is more common than in mill/moderate cases. TT3 and TSH levels are lower in COVID-19 patients than healthy individuals.

3.2. RESULTS OF META-ANALYSIS

3.2.1. TFT and COVID-19 Severity: Meta-Analysis 1 - TSH (µIU/mL) vs COVID-19 Severity

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exists different statistical methods to determine heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 49.54, $p < 0.00001$, $I^2 = 84\%$, in forest plot confidence intervals are overlapping in some studies. So, heterogeneity takes place and random effect model should be preferred in order to eliminate the errors that emerge from heterogeneity. Due to $p < 0.05$, there exists statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 0.3, lower limit = -0.05 (95% CI) and upper limit = 0.65 (95% CI) for random effect model. Among the mild/moderate and severe/critical groups standardized mean TSH difference is 0.3 greater in mild/moderate group within the 95% CI (-0.05 to 0.65).

Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. So, no relative size exists between the TSH values and COVID-19 severity. Table 3.1 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.1 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for random effect model.

Table 3.1. Model-based results for meta-analysis 1

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.34	0.21	0.48	49.54	8	0.000	84%	N/A	5.06	0.000
Random	0.3	-0.05	0.65	49.54	8	0.000	84%	0.23	1.7	0.09

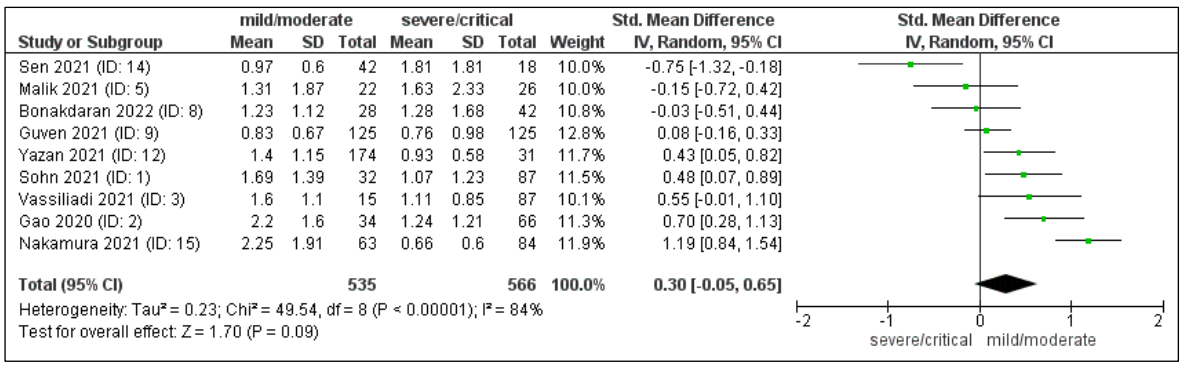

Figure 3.1. Random effect model result for meta-analysis 1

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.2. As can be seen from the figure, 77.7% of point (studies) locates inside the funnel shape and stays base of the graph which means studies are completed with small sample sizes. Distribution of points generates a symmetric figure. So, it can be interpreted as publication bias does not take place.

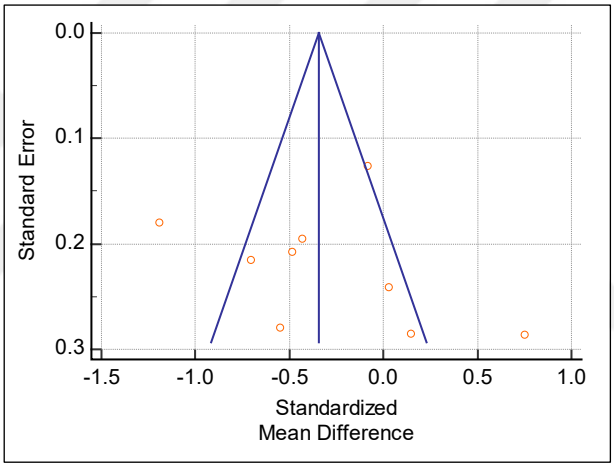

Figure 3.2. Funnel plot for meta-analysis 1

Egger's and Begg's test results are served in table 3.2. According to Egger's test, the p-values greater than 0.05, lead to symmetric distribution in the funnel plot. Since, p = 0.6799, funnel plot is in symmetric form and publication bias does not exists. According to Begg's test, p = 0.2971, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.2971$), so publication bias is not observed.

Table 3.2. Egger's & Begg's test results for meta-analysis 1

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
1.4158	-6.3650 to 9.1965	P = 0.6799	0.2778	P = 0.2971

3.2.2. TFT and COVID-19 Severity: Meta-Analysis 2 - FT3 (pg/ml) vs COVID-19 Severity

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As it mentioned in the introduction section, there exists different statistical methods to determine the heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75$ and in forest plot not overlapping the confidence intervals indicate heterogeneity.

In this meta-analysis Q-value = 16.19, $p = 0.003$, $I^2 = 75\%$, in forest plot confidence intervals are overlapping in some studies. So, heterogeneity takes place and random effect model should be preferred in order to eliminate the errors that emerge from heterogeneity. Due to $p < 0.05$, there exists statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 0.65, lower limit = 0.31 (95% CI) and upper limit = 0.99 (95% CI) for random effect model. Among the mild/moderate and severe/critical groups standardized mean FT3 difference is 0.65 greater in mild/moderate group within the 95% CI (0.31 to 0.99).

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,5 and 0,8 is evaluated as medium relative size. Table 3.3 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.3 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for random effect model.

Table 3.3. Model-based results for meta-analysis 2

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.72	0.56	0.88	16.19	4	0.003	75%	N/A	8.81	0.000
Random	0.65	0.31	0.99	16.19	4	0.003	75%	0.11	3.75	0.000

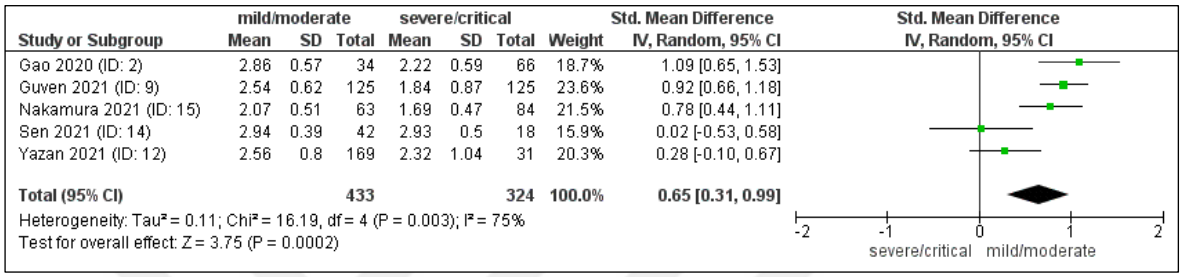

Figure 3.3. Random effect model result for meta-analysis 2

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.4. As can be seen from the figure, 60% of point (studies) locates inside the funnel shape and stays middle-base of the graph which means studies are completed with medium-small sample sizes. Distribution of points generates a symmetric figure. So, it can be interpreted as publication bias does not take place.

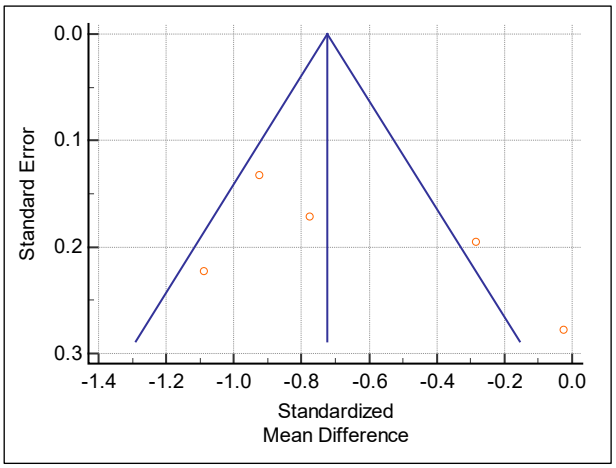

Figure 3.4. Funnel plot for meta-analysis 2

Egger's and Begg's test results are served in table 3.4. According to Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, p = 0.2962, funnel plot is in symmetric form and publication bias does not exists. According to the Begg's test, p = 0.1416, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.1416$), so publication bias is not observed.

Table 3.4. Egger's & Begg's test results for meta-analysis 2

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
4.3905	-6.6842 to 15.4653	P = 0.2962	0.6	P = 0.1416

3.2.3. TFT and COVID-19 Severity: Meta-Analysis 3 - TT3 (ng/dL) vs COVID-19 Severity

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exists different statistical methods to determine heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 45.81, $p < 0.00001$, $I^2 = 91\%$, in forest plot confidence intervals are overlapping in some studies. So, heterogeneity takes place and random effect model should be preferred in order to eliminate the errors that emerge from heterogeneity. Due to $p < 0.05$, there exists statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 1.19, lower limit = 0.36 (95% CI) and upper limit = 2.03 (95% CI) for random effect model. Among the mild/moderate and severe/critical groups standardized mean TT3 difference is 1.19 greater in mild/moderate group within the 95% CI (0.36 to 2.03).

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is greater or equal than 0,8 is evaluated as a large relative size. Table 3.5. presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.5 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for random effect model.

Table 3.5. Model-based results for meta-analysis 3

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.96	0.72	1.2	45.81	4	0.000	91%	N/A	7.83	0.000
Random	1.19	0.36	2.03	45.81	4	0.000	91%	0.82	2.8	0.005

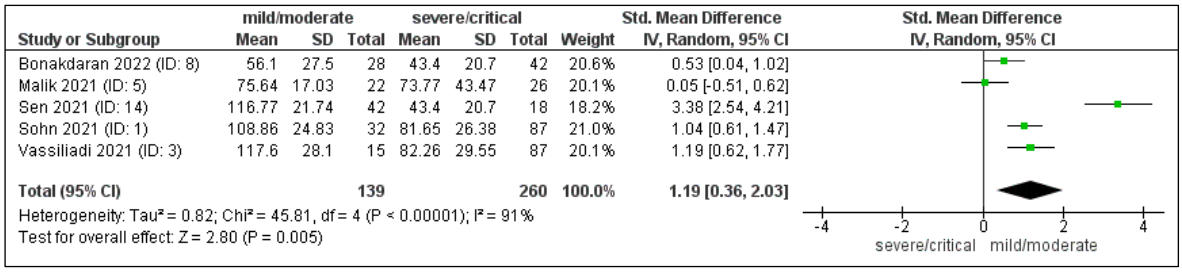

Figure 3.5. Random effect model result for meta-analysis 3

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.6. As can be seen from the figure, 40% of point (studies) locates inside the funnel shape and stays base of the graph which means studies are completed with small sample sizes. Distribution of points does not generate a symmetric figure. So, it can be interpreted as publication bias take place. But Egger's and Begg's should be done for the final bias existsence decision.

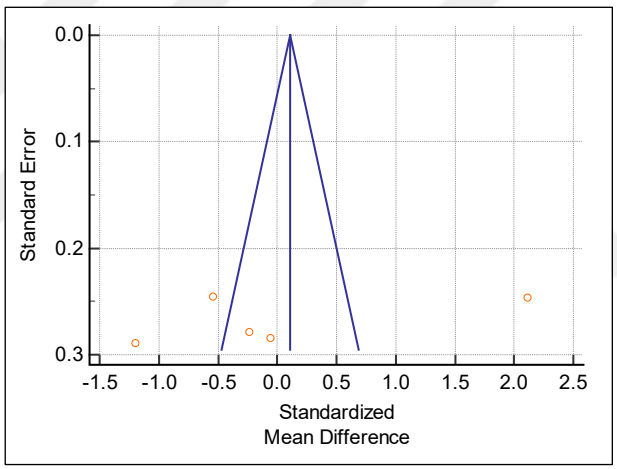

Figure 3.6. Funnel plot for meta-analysis 3

Egger's and Begg's test results are served in table 3.6. According to Egger's test, the p-values greater than 0.05, lead symmetric distribution in funnel plot. Since, $p = 0.3342$, funnel plot is in symmetric form and publication bias do not exists. According to the Begg's test, $p = 0.1416$, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.1416$), so publication bias is not observed.

Table 3.6. Egger's & Begg's test results for meta-analysis 3

Egger's Test			Begg's Test	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
-33.2515	-125.42 to 58.9169	P = 0.3342	-0.2000	P = 0.6242

3.2.4. TFT and COVID-19 Severity: Meta-Analysis 4 - FT4 (ng/dL) vs COVID-19 Severity

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As it mentioned in the introduction section, there exists different statistical methods to determine the heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 24.97, $p = 0.0003$, $I^2 = 76\%$, in forest plot confidence intervals are overlapping in some studies. So, heterogeneity take place and random effect model should be preferred in order to eliminate the errors emerge from heterogeneity. Due to $p < 0.05$, there exists statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 0.06, lower limit = -0.25 (95% CI) and upper limit = 0.37 (95% CI) for random effect model. Among the mild/moderate and severe/critical groups standardized mean FT4 difference is 0.06 greater in mild/moderate group within the 95% CI (-0.25 to 0.37).

Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. So, no relative size exists between the FT4 values and COVID-19 severity. Table 3.7. presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.7

shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for random effect model.

Table 3.7. Model-based results for meta-analysis 4

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.07	-0.07	0.22	24.97	6	0.000	76%	N/A	1.02	0.31
Random	0.06	-0.25	0.37	24.97	6	0.000	76%	0.13	0.4	0.69

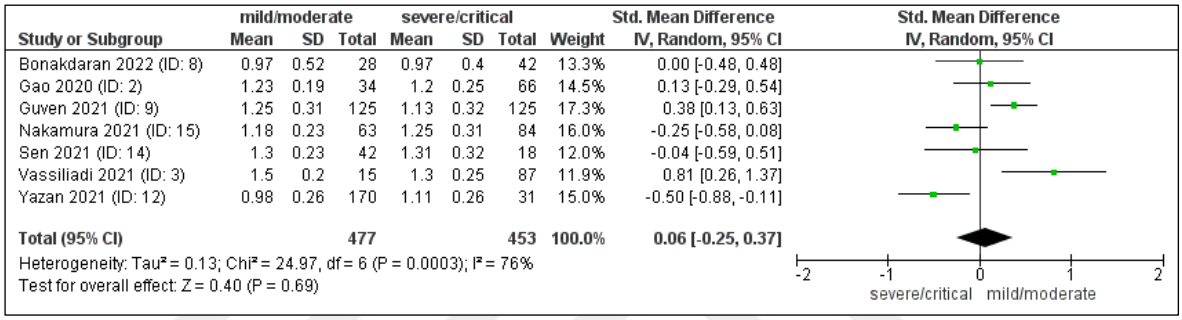

Figure 3.7. Random effect model result for meta-analysis 4

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.8. As can be seen from the figure, 57.14% of point (studies) locates inside the funnel shape and stays middle-base of the graph which means studies are completed with medium-small sample sizes. Distribution of points generate a symmetric figure. So, it can be interpreted as publication bias do not take place.

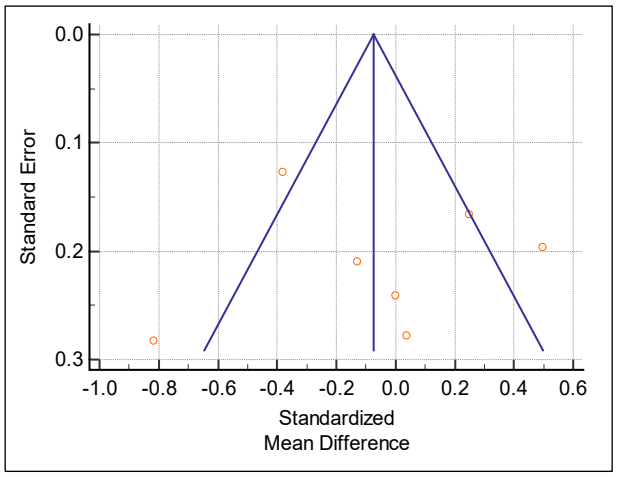

Figure 3.8. Funnel plot for meta-analysis 4

Egger's and Begg's test results are served in table 3.8. According to the Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, p = 0.8566, funnel plot is in symmetric form and publication bias do not exists. According to Begg's test, p = 0.8806, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.8806$), so publication bias is not observed.

Table 3.8. Egger's & Begg's test results for meta-analysis 4

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
0.5892	-7.3724 to 8.5508	P = 0.8566	-0.04762	P = 0.8806

3.2.5. TFT and COVID-19 Severity: Meta-Analysis 5 - TT4 (mcg/dL) vs COVID-19 Severity

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exists different statistical methods to determine heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 0.48, p = 0.49 , I^2 = 0%, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to p > 0.05, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 0.02, lower limit = -0.38 (95% CI) and upper limit = 0.42 (95% CI) for random effect model. Among the mild/moderate and severe/critical groups standardized mean TT4 difference is 0.02 greater in mild/moderate group within the 95% CI (-0.38 to 0.42).

Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. So, no relative size exists between the TT4 values and COVID-19 severity. Table 3.9 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.9 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for fixed effect model.

Table 3.9. Model-based results for meta-analysis 5

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.02	-0.38	0.42	0.48	1	0.49	0%	N/A	0.1	0.92
Random	0.02	-0.38	0.42	0.48	1	0.49	0%	0	0.1	0.92

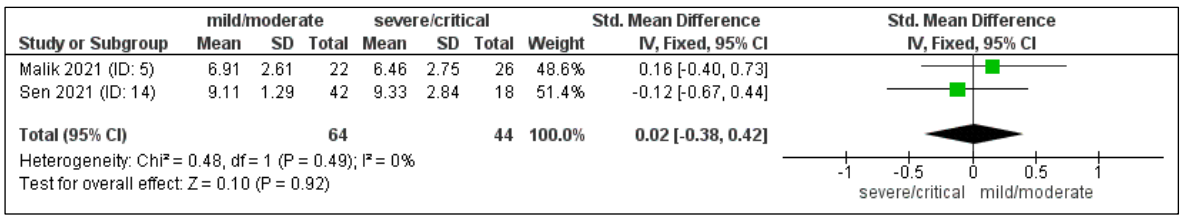

Figure 3.9. Fixed effect model result for meta-analysis 5

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.10. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays base of the graph which means studies are completed with small sample sizes. Distribution of points generates a symmetric figure. So, it can be interpreted as publication bias does not take place.

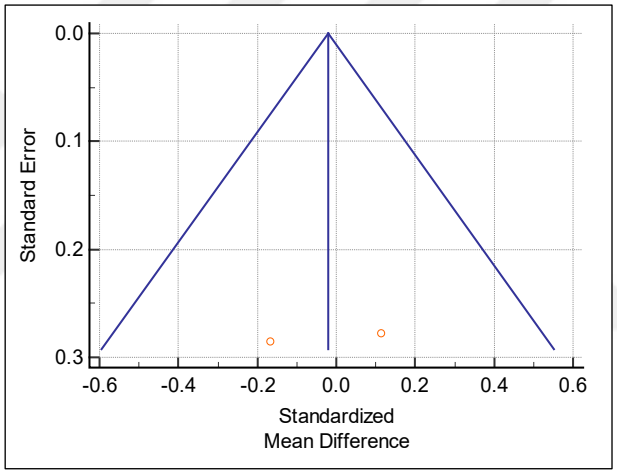

Figure 3.10. Funnel plot for meta-analysis 5

Egger's and Begg's test results are served in table 3.10. According to the Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, $p < 0.0001$, funnel plot is not in symmetric form and publication bias exists. Despite that, the view of funnel plot is symmetric. According to Begg's test, $p = 0.3173$, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.3173$), so publication bias is not observed.

Table 3.10. Egger's & Begg's test results for meta-analysis 5

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
-39.0760	N/A	P < 0.0001	-1	P = 0.3173

3.2.6. Comorbidity and COVID-19 Severity: Meta-Analysis 6 - Comorbidity (Diabetes) vs COVID-19 Severity

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exists different statistical methods to determine heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 3.16, $p = 0.08$, $I^2 = 68\%$, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to $p > 0.05$, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using the odds ratio. Standardized mean difference (SMD), can not be calculated because data is not in continuous form. ES = 0.46, lower limit = 0.25 (95% CI) and upper limit = 0.86 (95% CI) for random effect model. Among the mild/moderate and severe/critical groups odds ratio for diabetes is 0.46 greater in mild/moderate group within the 95% CI (0.25 to 0.86) which means the prognosis of COVID-19 is milder in patients having diabetes.

According to Cohen' (1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. Table 3.11 presents the odds ratio, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.11 shows the study-based mean, SD, weight, odds ratio, 95% CI, forest plot for random fixed model.

Table 3.11. Model-based results for meta-analysis 6

Model	OR	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.46	0.25	0.86	3.16	1	0.08	68%	N/A	2.41	0.02
Random	0.54	0.16	1.79	3.16	1	0.08	68%	0.52	1.01	0.31

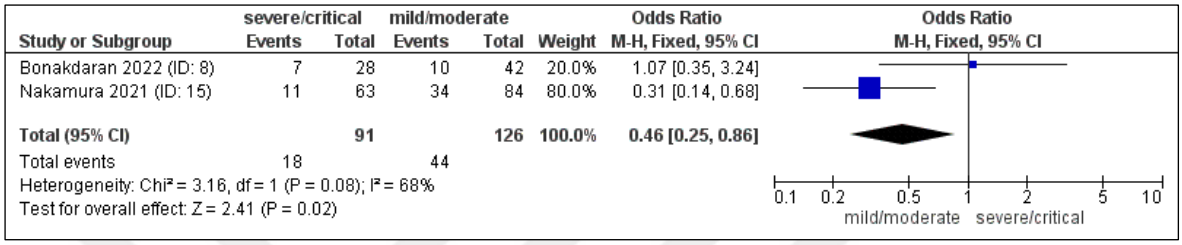

Figure 3.11. Fixed effect model result for meta-analysis 6

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.12. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and middle-stays base of the graph which means studies are completed with medium-small sample sizes. Distribution of points does not generate a symmetric figure. So, it can be interpreted as publication bias take place.

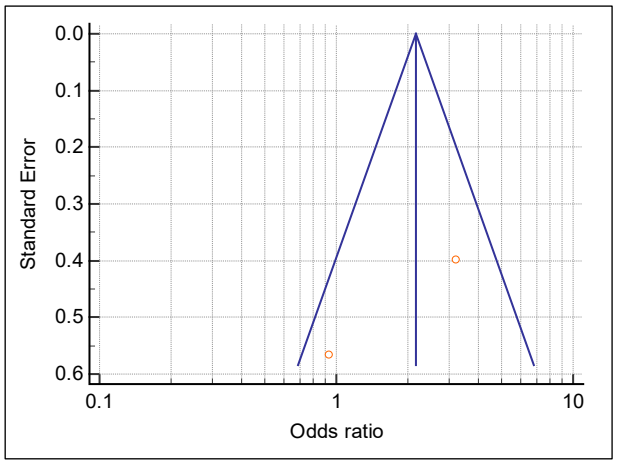

Figure 3.12. Funnel plot for meta-analysis 6

Egger's and Begg's test results are served in table 3.12. According to the Egger's test, the p-values greater than 0.05, lead symmetric distribution in funnel plot. Since, $p < 0.0001$, funnel plot is not in symmetric form and publication bias exists. According to Begg's test, $p = 0.3173$, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.3173$), so publication bias is not observed.

Table 3.12. Egger's & Begg's test results for meta-analysis 6

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
-7.3447	N/A	P < 0.0001	-1	P = 0.3173

3.2.7. Comorbidity and COVID-19 Severity: Meta-Analysis 7 - Comorbidity (Hypertension) vs COVID-19 Severity

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exists different statistical methods to determine heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 0.27, p = 0.6, I^2 = 0%, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to p > 0.05, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using the odds ratio. Standardized mean difference (SMD), can not be calculated because data is not in continuous form. ES = 0.36, lower limit = 0.2 (95% CI) and upper limit = 0.68 (95% CI) for random effect model. Among the mild/moderate and severe/critical groups odds ratio for hypertension is 0.36 greater in mild/moderate group within the 95% CI (0.2 to 0.68) which means the prognosis of COVID-19 is milder in patients having hypertension.

According to Cohen' (1988) "Rules-of-Thumb" for the standardized mean difference of the effect size is between 0,2 and 0,5 is evaluated as small relative size. Table 3.13 presents the odds ratio, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.13 shows the study-based mean, SD, weight, odds ratio, 95% CI, forest plot for fixed effect model.

Table 3.13. Model-based results for meta-analysis 7

Model	OR	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.36	0.19	0.67	0.27	1	0.6	0%	N/A	3.25	0.001
Random	0.36	0.2	0.68	0.27	1	0.6	0%	0	3.19	0.001

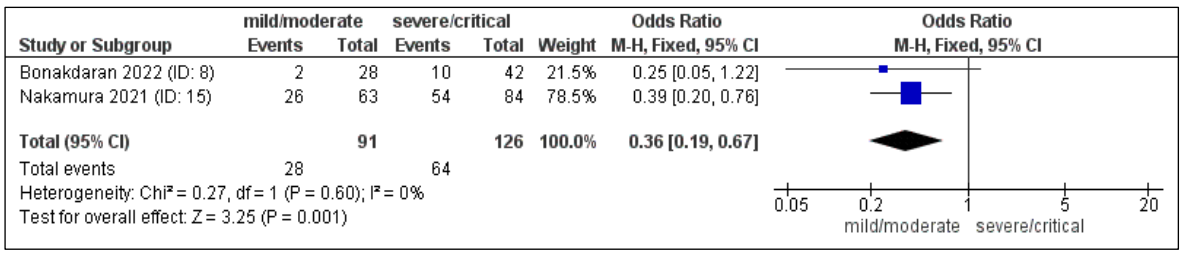

Figure 3.13. Fixed effect model result for meta-analysis 7

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.14. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays middle-base of the graph which means studies are completed with medium-small sample sizes. Distribution of points does not generate a symmetric figure. So, it can be interpreted as publication bias take place.

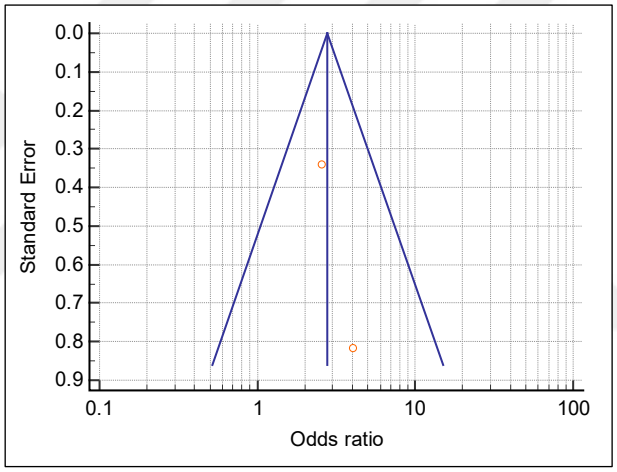

Figure 3.14. Funnel plot for meta-analysis 7

Egger's and Begg's test results are served in table 3.14. According to the Egger's test, the p-values greater than 0.05, lead symmetric distribution in funnel plot. Since, $p < 0.0001$, funnel plot is not in symmetric form and publication bias exists. According to Begg's test, $p = 0.3173$, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.3173$), so publication bias is not observed.

Table 3.14. Egger's & Begg's test results for meta-analysis 7

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
0.9693	N/A	P < 0.0001	1	P = 0.3173

3.2.8. TFT and Mortality Rates by COVID-19: Meta-Analysis 8 - TSH (µIU/mL) vs Mortality

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exists different statistical methods to determine the heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 2.04, $p = 0.73$, $I^2 = 0\%$, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to $p > 0.05$, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 0.26, lower limit = 0.05 (95% CI) and upper limit = 0.47 (95% CI) for random effect model. Among the survived and deceased groups standardized mean TSH difference is 0.26 greater in the survived group within the 95% CI (0.05 to 0.47).

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. Table 3.15 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.15 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for fixed effect model.

Table 3.15. Model-based results for meta-analysis 8

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.26	0.05	0.47	2.04	4	0.73	0%	N/A	2.48	0.01
Random	0.26	0.05	0.47	2.04	4	0.73	0%	0	2.48	0.01

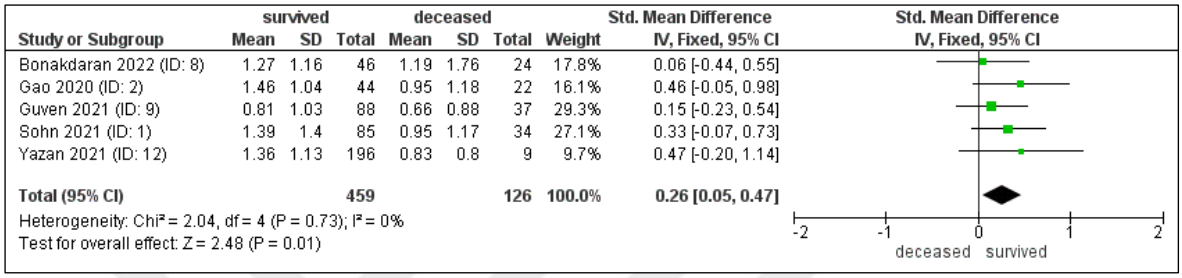

Figure 3.15. Fixed effect model result for meta-analysis 8

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.16. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays middle of the graph which means studies are completed with medium sample sizes. Distribution of points generates a symmetric figure. So, it can be interpreted as publication bias does not take place.

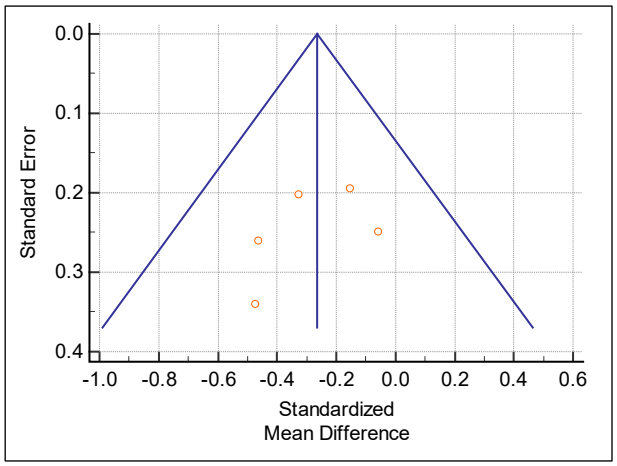

Figure 3.16. Funnel plot for meta-analysis 8

Egger's and Begg's test results are served in table 3.16. According to Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, p = 0.4159, funnel plot is in symmetric form and publication bias does not exist. According to Begg's test, p = 0. 3271, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.3271$), so publication bias is not observed.

Table 3.16. Egger's & Begg's test results for meta-analysis 8

Egger's Test			Begg's Test	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
-1.6437	-7.1992 to 3.9119	P = 0.4159	-0.4	P = 0.3272

3.2.9. TFT and Mortality Rates by COVID-19: Meta-Analysis 9 - FT3 (pg/mL) vs Mortality

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exist different statistical methods to determine heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 4.31, p = 0.77, I^2 = 54%, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to p > 0.05, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 0.77, lower limit = 0.49 (95% CI) and upper limit = 1.06 (95% CI) for random effect model. Among the survived and deceased groups standardized mean FT3 difference is 0.77 greater in survived group within the 95% CI (0.49 to 1.06).

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,5 and 0,8 is evaluated as medium relative size. Table 3.17 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.17 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for fixed effect model.

Table 3.17. Model-based results for meta-analysis 9

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.77	0.49	1.06	4.31	2	0.12	54%	N/A	5.26	0.000
Random	0.83	0.38	1.27	4.31	2	0.12	54%	0.08	3.6	0.000

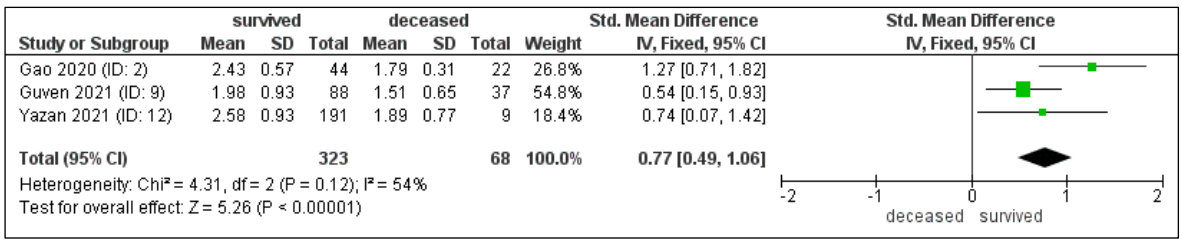

Figure 3.17. Fixed effect model result for meta-analysis 9

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.18. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays middle of the graph which means studies are completed with medium sample sizes. Distribution of points generates a symmetric figure. So, it can be interpreted as publication bias does not take place.

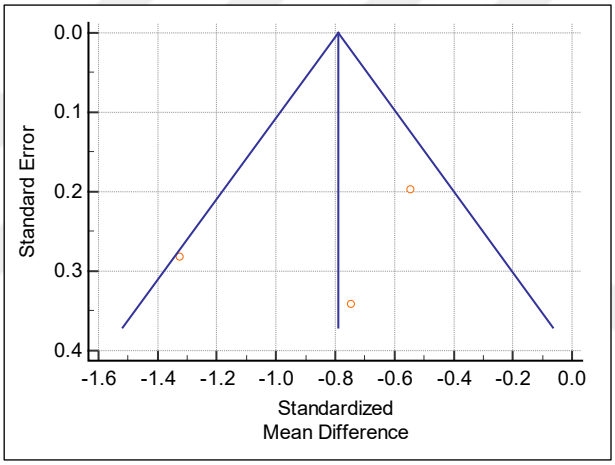

Figure 3.18. Funnel plot for meta-analysis 9

Egger's and Begg's test results are served in table 3.18. According to the Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, p = 0.6125, funnel plot is in symmetric form and publication bias does not exists. According to Begg's test, p = 0.6015, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.6015$), so publication bias is not observed.

Table 3.18. Egger's & Begg's test results for meta-analysis 9

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
-3.2927	-63.3271 to 56.7417	P = 0.6125	-0.3333	P = 0.6015

3.2.10. TFT and Mortality Rates by COVID-19: Meta-Analysis 10 - TT3 (ng/dL) vs Mortality

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exist different statistical methods to determine heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 1.16, $p = 0.28$, $I^2 = 54\%$, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to $p > 0.05$, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 0.73, lower limit = 0.41 (95% CI) and upper limit = 1.05 (95% CI) for random effect model. Among the survived and deceased groups standardized mean TT3 difference is 0.73 greater in survived group within the 95% CI (0.41 to 1.05).

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,5 and 0,8 is evaluated as medium relative size. Table 3.19 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.19 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for fixed effect model.

Table 3.19. Model-based results for meta-analysis 10

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.73	0.41	1.05	1.16	1	0.28	14%	N/A	4.47	0.000
Random	0.72	0.38	1.07	1.16	1	0.28	14%	0.01	4.1	0.000

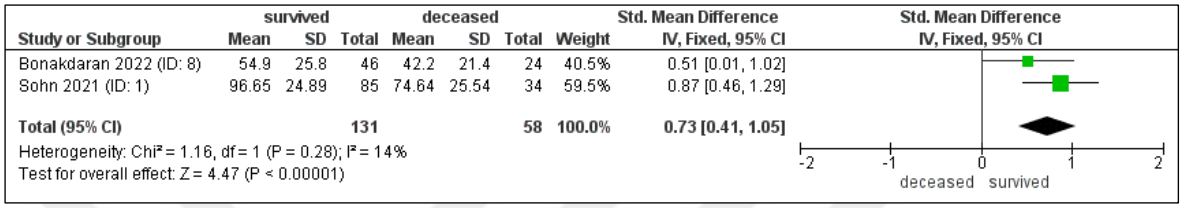

Figure 3.19. Fixed effect model result for meta-analysis 10

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.20. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays middle of the graph which means studies are completed with medium sample sizes. Distribution of points generates a symmetric figure. So, it can be interpreted as publication bias does not take place.

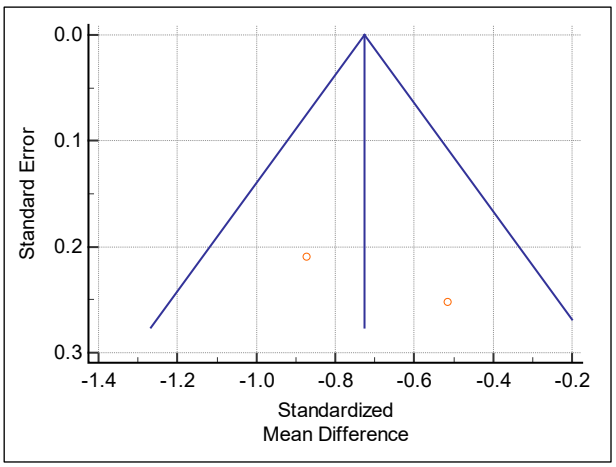

Figure 3.20. Funnel plot for meta-analysis 10

Egger's and Begg's test results are served in table 3.20. According to Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, $p < 0.0001$, funnel plot is not in symmetric form and publication bias exists. Despite that, the view of funnel plot is symmetric. According to Begg's test, $p = 0.3173$, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.3173$), so publication bias is not observed.

Table 3.20. Egger's & Begg's test results for meta-analysis 10

Egger's Test			Begg's Test	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
8.2362	N/A	P < 0.0001	1	P = 0.3173

3.2.11. TFT and Mortality Rates by COVID-19: Meta-Analysis 11 - FT4 (ng/dL) vs Mortality

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As it mentioned in the introduction section, there exist different statistical methods to determine the heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 5.15, $p = 0.27$, $I^2 = 22\%$, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to $p > 0.05$, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 0.13, lower limit = -0.07 (95% CI) and upper limit = 0.33 (95% CI) for random effect model. Among the survived and deceased groups standardized mean FT4 difference is 0.13 greater in survived group within the 95% CI (-0.07 to 0.33).

Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. So, no relative size exists between the FT4 and COVID-19 mortality.Table 3.21 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.21 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for fixed effect model.

Table 3.21. Model-based results for meta-analysis 11

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.13	-0.07	0.33	5.15	4	0.27	22%	N/A	1.3	0.19
Random	0.12	-0.1	0.35	5.15	4	0.27	22%	0.02	1.07	0.28

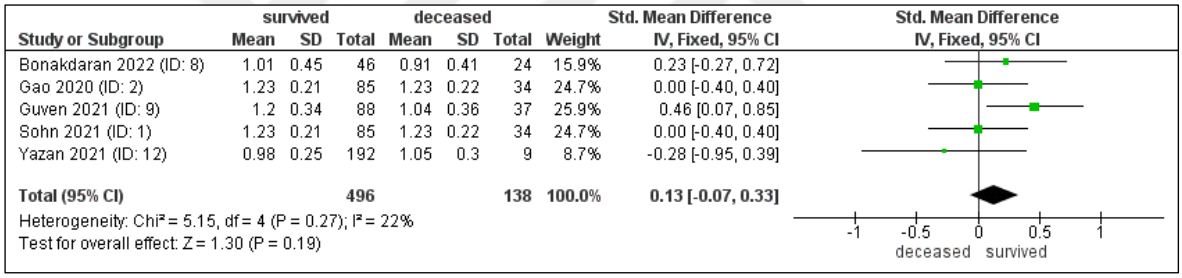

Figure 3.21. Fixed effect model result for meta-analysis 11

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.22. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays middle of the graph which means studies are completed with medium sample sizes. Distribution of points generates a symmetric figure. So, it can be interpreted as publication bias does not take place.

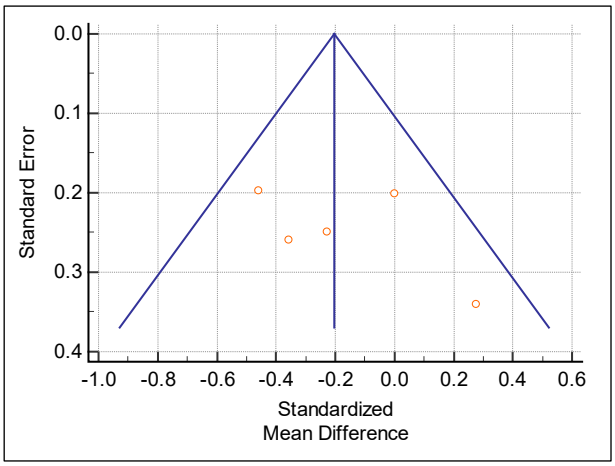

Figure 3.22. Funnel plot for meta-analysis 11

Egger's and Begg's test results are served in table 3.22. According to Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, p = 0.3994, funnel plot is in symmetric form and publication bias does not exists. According to Begg's test, p = 0.3272, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.3272$), so publication bias is not observed.

Table 3.22. Egger's & Begg's test results for meta-analysis 11

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
2.6688	-5.9992 to 11.3368	P = 0.3994	0.4	P = 0.3272

3.2.12. Thyroid Dysfunction and COVID-19 Severity: Meta-Analysis 12 - Thyroid Dysfunction vs COVID-19 Severity

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As it mentioned in the introduction section, there exist different statistical methods to determine the heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 9.74, p = 0.14, I^2 = 38%, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to p > 0.05, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using the odds ratio. Standardized mean difference (SMD), can not be calculated because data is not in continuous form. ES = 4.8, lower limit = 3.03 (95% CI) and upper limit = 7.6 (95% CI) for random effect model. Among the mild/moderate and severe/critical groups odds ratio for having thyroid dysfunction is 4.8 greater in severe/critical group within the 95% CI (3.03 to 7.6) which means the prognosis of COVID-19 is more severe in patients having thyroid dysfunction.

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is greater or equal than 0,8 is evaluated as a large relative size. Table 3.23 presents the odds ratio, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.23 shows the study-based mean, SD, weight, odds ratio, 95% CI, forest plot for fixed effect model.

Table 3.23. Model-based results for meta-analysis 12

Model	OR	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	4.8	3.03	7.6	9.74	6	0.14	38%	N/A	6.68	0.0000
Random	4.91	2.61	9.21	9.74	6	0.14	38%	0.27	4.94	0.0000

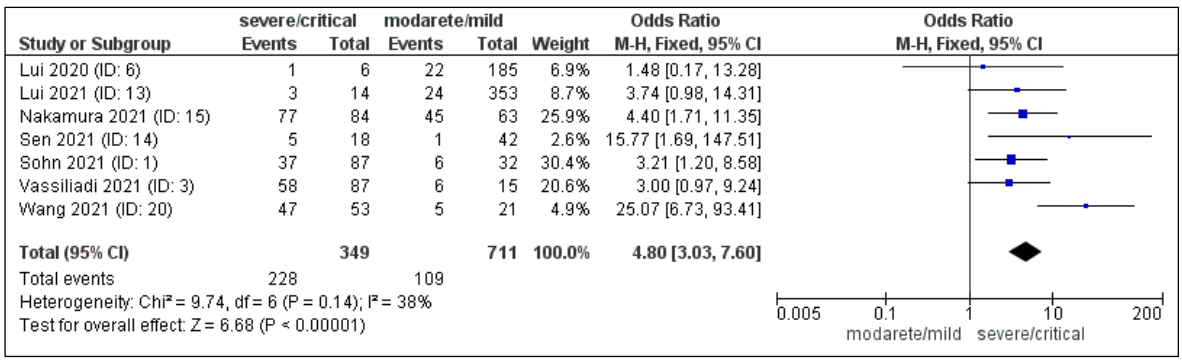

Figure 3.23. Fixed effect model result for meta-analysis 12

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. Ifthe view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.24. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays base to the upper-middle of the graph which means studies are completed with medium-small sample sizes. Distribution of points generates a symmetric figure. So, it can be interpreted as publication bias does not take place.

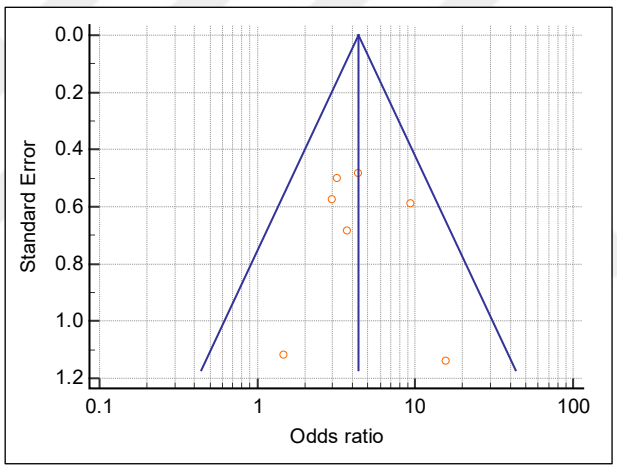

Figure 3.24. Funnel plot for meta-analysis 12

Egger's and Begg's test results are served in table 3.24. According to the Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, p = 0.14, funnel plot is in symmetric form and publication bias does not exists. According to Begg's test, p = 0.815, and the statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant (0.05 < 0. 815), so publication bias is not observed.

Table 3.24. Egger's & Begg's test results for meta-analysis 12

Egger's Test			Begg's Test	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
0.3212	-3.0254 to 3.6678	P = 0.8150	-0.04762	P = 0.8806

3.2.13. TFT and COVID-19 Existence: Meta-Analysis 13 - TSH (μIU/mL) vs COVID-19 Existence

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As it mentioned in the introduction section, there exist different statistical methods to determine the heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 2.73, $p = 0.25$, $I^2 = 27\%$, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to $p > 0.05$, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = -0.39, lower limit = -0.58 (95% CI) and upper limit = -0.2 (95% CI) for random effect model. Among the COVID-19 and non-COVID-19 patients groups, standardized mean TSH difference is 0.39 greater in non-COVID-19 group within the 95% CI (-0.58 to -0.2).

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. Table 3.25 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.25 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for fixed effect model.

Table 3.25. Model-based results for meta-analysis 13

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	-0.39	-0.58	-0.2	2.73	2	0.25	27%	N/A	4.01	0.000
Random	-0.38	-0.6	-0.15	2.73	2	0.25	27%	0.01	2.37	0.001

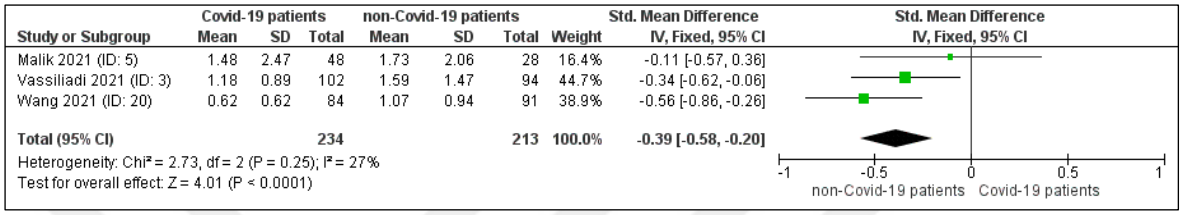

Figure 3.25. Fixed effect model result for meta-analysis 13

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.26. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays middle of the graph which means studies are completed with medium sample sizes. Distribution of points generate a symmetric figure. So, it can be interpreted as publication bias does not take place.

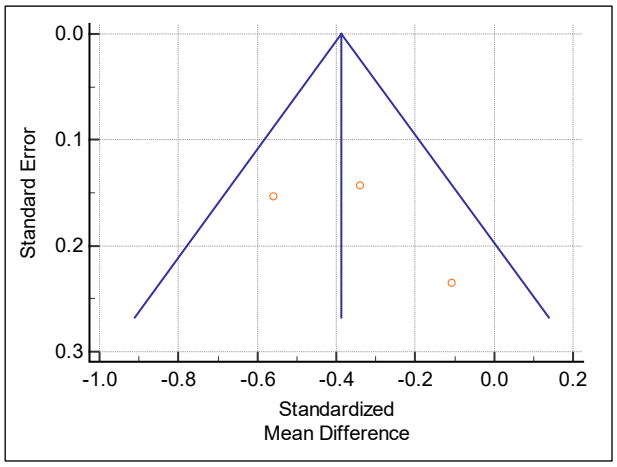

Figure 3.26. Funnel plot for meta-analysis 13

Egger's and Begg's test results are served in table 3.26. According to the Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, p = 0.52, funnel plot is in symmetric form and publication bias does not exists. According to Begg's test, p = 0.6015, and statistically significant level is evaluated as p < 0.05. P-value for this test is not statistically significant (0.05 < 0. 6015), so publication bias is not observed.

Table 3.26. Egger's & Begg's test results for meta-analysis 13

Egger's Test			Begg's Test	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
3.3312	-41.7435 to 48.4059	P = 0.52	0.3333	P = 0.6015

3.2.14. TFT and COVID-19 Existence: Meta-Analysis 14 - TT3 (ng/dL) vs COVID-19 Existence

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exist different statistical methods to determine heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 11.57, p = 0.003, I^2 = 83%, in forest plot confidence intervals are overlapping in some studies. So, heterogeneity takes place and random effect model should be preferred in order to eliminate the errors that emerge from heterogeneity. Due to $p < 0.05$, there exists statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = -0.02, lower limit = -0.49 (95% CI) and upper limit = 0.45 (95% CI) for random effect model. Among the COVID-19 and non-COVID-19 patients groups, standardized mean TT3 difference is 0.02 greater in non-COVID-19 group within the 95% CI (-0.49 to 0.45).

Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. So, no relative size exists between the TT3 values and COVID-19 existence. Table 3.27 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.27 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for random effect model.

Table 3.27. Model-based results for meta-analysis 14

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.1	-0.09	0.29	11.57	2	0.003	83%	N/A	1.02	0.31
Random	-0.02	-0.49	0.45	11.57	2	0.003	83%	0.14	0.09	0.93

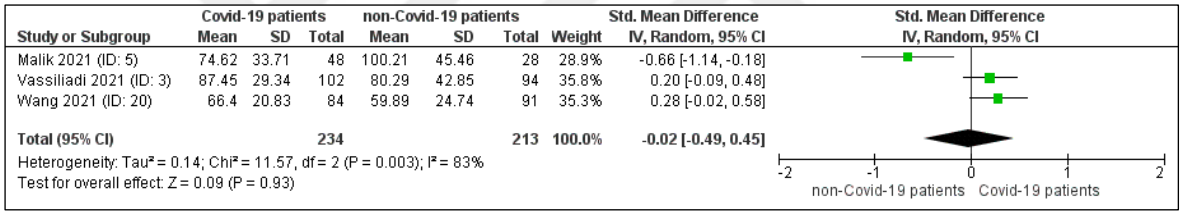

Figure 3.27. Random effect model result for meta-analysis 14

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.28. As can be seen from the figure, 66.67% of point (studies) locates inside the funnel shape and stays middle of the graph which means studies are completed with medium sample sizes. Distribution of points generates a asymmetric figure. So, it can be interpreted as publication bias take place.

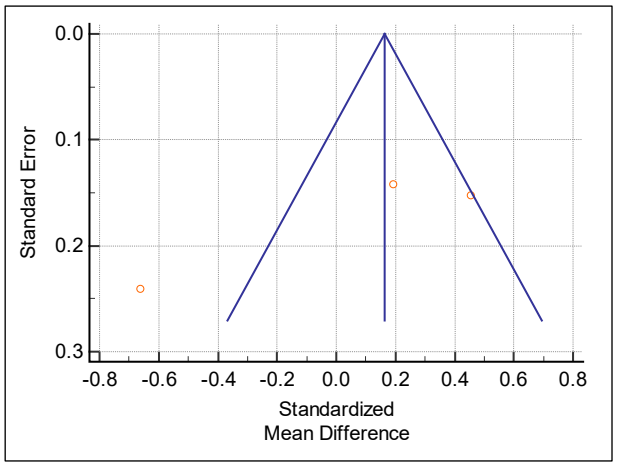

Figure 3.28. Funnel plot for meta-analysis 14

Egger's and Begg's test results are served in table 3.28. According to the Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, p = 0.2894, funnel plot is in symmetric form and publication bias does not exists. According to Begg's test, p = 0.6015, and statistically significant level is evaluated as p < 0.05. P-value for this test is not statistically significant (0.05 < 0. 6015), so publication bias is not observed.

Table 3.28. Egger's & Begg's test results for meta-analysis 14

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
-9.7585	-70.3579 to 50.8409	P = 0.2894	-0.3333	P = 0.6015

3.2.15. TFT and COVID-19 Existence: Meta-Analysis 15 - TT4 (mcg/dL) vs COVID-19 Existence

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As it mentioned in the introduction section, there exist different statistical methods to determine the heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If p < 0.05, I^2 > 75% and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 3.67, p = 0.06, I^2 = 73%, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to p > 0.05, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = 0.47, lower limit = 0.22 (95% CI) and upper limit = 0.73 (95% CI) for random effect model. Among the COVID-19 and non-COVID-19 patients groups standardized mean TT4 difference is 0.47 greater in COVID-19 group within the 95% CI (0.22 to 0.73).

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. Table 3.29 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.29 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for fixed effect model.

Table 3.29. Model-based results for meta-analysis 15

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	0.47	0.22	0.73	3.67	1	0.06	73%	N/A	3.63	0.000
Random	0.39	-0.14	0.92	3.67	1	0.06	73%	0.11	1.45	0.15

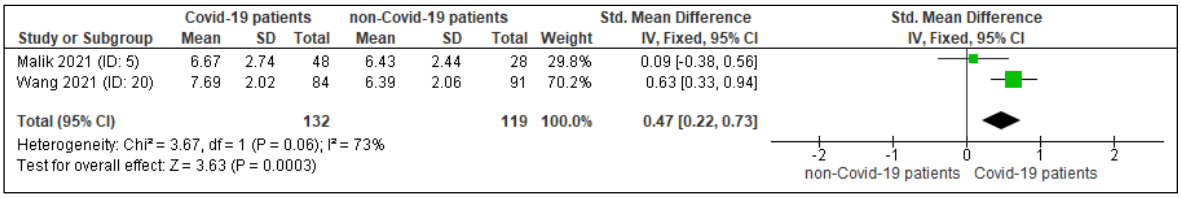

Figure 3.29. Fixed effect model result for meta-analysis 15

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study.

Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.30. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays middle of the graph which means studies are completed with medium sample sizes. Distribution of points generates a symmetric figure. So, it can be interpreted as publication bias does not take place.

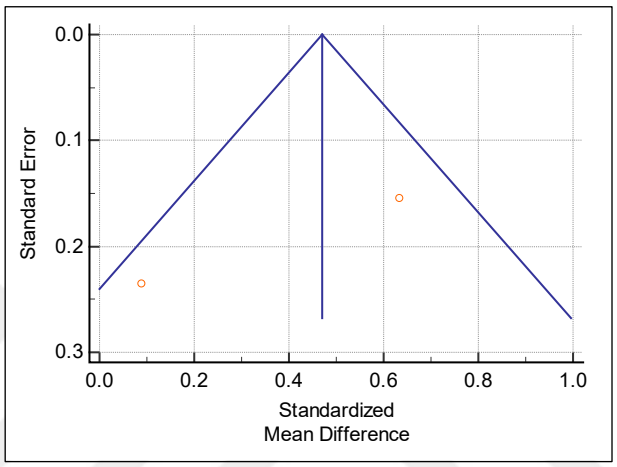

Figure 3.30. Funnel plot for meta-analysis 15

Egger's and Begg's test results are served in table 3.30. According to the Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, $p < 0.0001$, funnel plot is not in symmetric form and publication bias exists. According to Begg's test, $p = 0.3173$, and the statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.3173$), so publication bias is not observed.

Table 3.30. Egger's & Begg's test results for meta-analysis 15

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
-6.7103	-70.3579 to 50.8409	P < 0.0001	-1	P = 0.3173

3.2.16. TFT and COVID-19 Existence: Meta-Analysis 16 - TSH (μIU/mL) vs COVID-19/Healthy Status

Heterogeneity Tests, Statistical Model Selection and Effect Size

Homogeneity and heterogeneity of the studies lead to determining the type of model. Basically, heterogeneity shows the variety among the studies. As mentioned in the introduction section, there exist different statistical methods to determine heterogeneity. Q-value –measured as Chi-squared-, I^2 value -tells the amount of heterogeneity- and forest plot is observed by using the Review Manager software. If $p < 0.05$, $I^2 > 75\%$ and in forest plot not overlapping the confidence intervals indicates heterogeneity.

In this meta-analysis Q-value = 2.53, $p = 0.11$, $I^2 = 61\%$, in forest plot confidence intervals are overlapping. So, heterogeneity does not take place and fixed effect model is appropriate to use. Due to $p > 0.05$, there exists no statistically significant difference among the studies.

Effect size (ES) is measured by using standardized mean difference (SMD), since the data is continuous -not binary- odds ratio can not be calculated. ES = -1.08, lower limit = -1.3 (95% CI) and upper limit = -0.87 (95% CI) for random effect model. Among the COVID-19 patients and non-COVID-19 healthy groups, standardized mean TSH difference is 1.08 greater in COVID-19 patients group within the 95% CI (-1.3 to -0.87).

According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is greater than 0.8, is evaluated as high relative size. Table 3.31 presents the standardized mean difference, lower/upper limits, heterogeneity test results and test of null metrics based on both fixed and random effect models. Figure 3.31 shows the study-based mean, SD, weight, standardized mean difference, 95% CI, forest plot for fixed effect model.

Table 3.31. Model-based results for meta-analysis 16

Model	SMD	95% CI		Heterogeneity					Test of null (2-Tail)	
		Lower	Upper	Chi^2	df	p-value	I^2	Tau^2	z-value	p-value
Fixed	-1.08	-1.3	-0.87	2.53	1	0.11	61%	N/A	9.92	0.000
Random	-1.2	-1.68	-0.79	2.53	1	0.11	61%	0.08	4.87	0.000

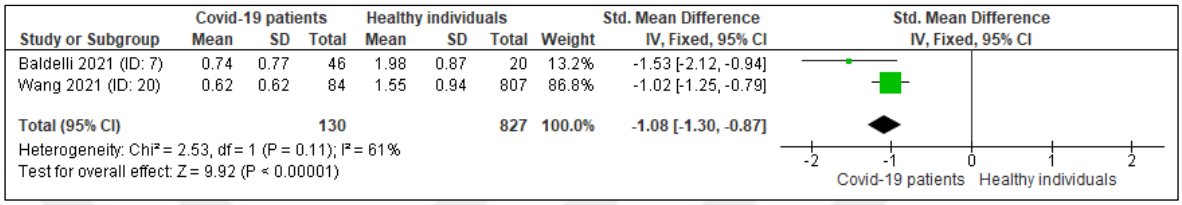

Figure 3.31. Fixed effect model result for meta-analysis 16

Publication Bias

For assessing the publication bias among the studies funnel plot is generated. If the view of the graph is like a symmetric funnel, it indicates that there is not any bias in the study. Since, it is not a quantitive way to interpret and may change person to person, rank correlation (Begg's test) and linear regression (Egger's test) tests are conducted.

Funnel plot is presented in figure 3.32. As can be seen from the figure, 100% of point (studies) locates inside the funnel shape and stays base and middle of the graph which means studies are completed with small/medium sample sizes. Distribution of points generate a symmetric figure. So, it can be interpreted as publication bias does not take place.

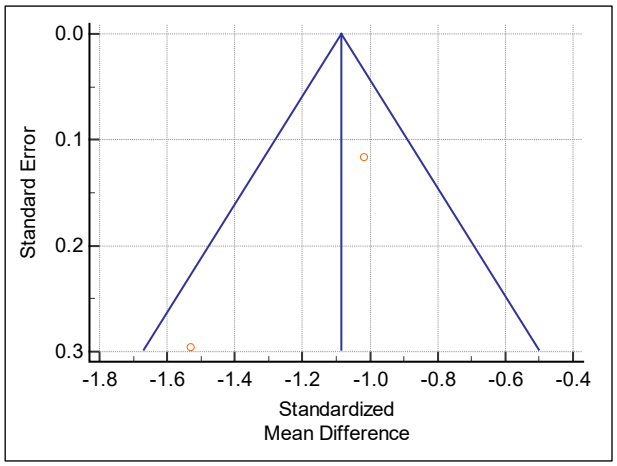

Figure 3.32. Funnel plot for meta-analysis 16

Egger's and Begg's test results are served in table 3.32. According to Egger's test, the p-values greater than 0.05, lead to symmetric distribution in funnel plot. Since, $p < 0.0001$, funnel plot is not in symmetric form and publication bias exists. According to Begg's test, $p = 0.3173$, and statistically significant level is evaluated as $p < 0.05$. P-value for this test is not statistically significant ($0.05 < 0.3173$), so publication bias is not observed.

Table 3.32. Egger's & Begg's test results for meta-analysis 16

Egger's Test			**Begg's Test**	
Intercept	**95% CI**	**Significant level**	**Kendall's Tau**	**Significant level**
-2.8686	N/A	P < 0.0001	-1	P = 0.3173

Effect size and relative size according to Cohen'(1988) for each meta-analysis are summarized in the table 3.33.

Table 3.33. Relative Size for meta-analyses

Meta-Analysis	Effect Size	Relative Size
Meta-Analysis 1: TSH (μIU/mL) vs COVID-19 Severity	0.3	no
Meta-Analysis 2: FT3 (pg/ml) vs COVID-19 Severity	0.65	medium
Meta-Analysis 3: TT3 (ng/dL) vs COVID-19 Severity	1.19	large
Meta-Analysis 4: FT4 (ng/dL) vs COVID-19 Severity	0.06	no
Meta-Analysis 5: TT4 (mcg/dL) vs COVID-19 Severity	0.02	no
Meta-Analysis 6: Comorbidity (Diabetes) vs COVID-19 Severity	0.46	small
Meta-Analysis 7: Comorbidity (Hypertension) vs COVID-19 Severity	0.36	small
Meta-Analysis 8: TSH (μIU/mL) vs Mortality	0.26	small
Meta-Analysis 9: FT3 (pg/mL) vs Mortality	0.77	medium
Meta-Analysis 10: TT3 (ng/dL) vs Mortality	0.73	medium
Meta-Analysis 11: FT4 (ng/dL) vs Mortality	0.13	no
Meta-Analysis 12: Thyroid Dysfunction vs COVID-19 Severity	4.8	large
Meta-Analysis 13: TSH (μIU/mL) vs COVID-19 Existence	-0.39	small
Meta-Analysis 14: TT3 (ng/dL) vs COVID-19 Existence	-0.02	no
Meta-Analysis 15: TT4 (mcg/dL) vs COVID-19 Existence	0.36	small
Meta-Analysis 16: TSH (μIU/mL) vs COVID-19/Healthy Status	-1.08	large

4. CONCLUSION, DISCUSSION AND LIMITATIONS

A meta-analysis was conducted on the effects of COVID-19 on thyroid disorders. Under five main research question, total 16 meta-analyses were completed. Among the 4260 reviewed articles, 13 of them met the inclusion criteria and were added to the meta-analyses. Correlation, heterogeneity, and publication bias tests were taken into consideration. For each meta-analyses conclusion are served below.

4.1. CONCLUSION

First research question is trying to discover the TFT and COVID-19 severity relationship. Focused on TSH, FT3, TT3, FT4, TT4 hormone tests.

TFT and COVID-19 Severity: Meta-Analysis 1: TSH (μIU/mL) vs COVID-19 Severity

Heterogeneity takes place ($Q = 49.54$, $p < 0.00001$, $I^2 = 84\%$), random effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 0.3 (-0.05 to 0.65). Among the mild/moderate and severe/critical groups standardized mean TSH difference is 0.3 greater in mild/moderate group within the 95% CI (-0.05 to 0.65). Due to $p < 0.05$, there exists statistically significant difference among the studies. Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. So, no relative size exists between the TSH values and COVID-19 severity. Publication bias is not observed regards to funnel plot, Egger and Begg tests.

TFT and COVID-19 Severity: Meta-Analysis 2: FT3 (pg/ml) vs COVID-19 Severity

Heterogeneity takes place ($Q = 16.19$, $p = 0.003$, $I^2 = 75\%$), random effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 0.65 (0.31 to 0.99). Among the mild/moderate and severe/critical groups standardized mean FT3 difference is 0.65 greater in mild/moderate group within the 95% CI (0.31 to 0.99). Due to $p < 0.05$, there exists statistically significant difference among the studies. According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,5 and 0,8 is evaluated as medium relative size. So, FT3 values

have medium effect on the COVID-19 severity. Publication bias is not observed regards to funnel plot, Egger and Begg tests.

TFT and COVID-19 Severity: Meta-Analysis 3: TT3 (ng/dL) vs COVID-19 Severity

Heterogeneity takes place (Q = 45.81, p < 0.00001, I^2 = 91%), random effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 1.19 (0.36 to 2.03). Among the mild/moderate and severe/critical groups standardized mean TT3 difference is 1.19 greater in mild/moderate group within the 95% CI (0.36 to 2.03). Due to p < 0.05, there exists statistically significant difference among the studies. According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is greater or equal than 0,8 is evaluated as a large relative size. So, TT3 values have large effect on the COVID-19 severity. Publication bias is observed regards to funnel plot but not observed regards to Egger and Begg tests.

TFT and COVID-19 Severity: Meta-Analysis 4: FT4 (ng/dL) vs COVID-19 Severity

Heterogeneity takes place (Q = 24.97, p = 0.0003, I^2 = 76%), random effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 0.06 (-0.25 to 0.37). Among the mild/moderate and severe/critical groups standardized mean FT4 difference is 0.06 greater in mild/moderate group within the 95% CI (-0.25 to 0.37). Due to p < 0.05, there exists statistically significant difference among the studies. Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. Publication bias is not observed regards to funnel plot, Egger and Begg tests.

TFT and COVID-19 Severity: Meta-Analysis 5: TT4 (mcg/dL) vs COVID-19 Severity

Heterogeneity do not takes place (Q = 0.48, p = 0.49 , I^2 = 0%), fixed effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 0.02 (-0.38 to 0.42). Among the mild/moderate and severe/critical groups standardized mean TT4 difference is 0.02 greater in mild/moderate group within the 95% CI (-0.38 to 0.42). Due to p > 0.05, there exists no statistically significant difference among the studies. Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. Publication bias is not observed regards to funnel plot and Begg test but observed regards to Egger test.

In conclusion, TT3 value has large; FT3 value has medium; TSH have small; FT4 and TT4 have very small or no effect on the COVID-19 severity. Among the TFT's, TT3 has the biggest and TT4 has the smallest effect on COVID-19 severity Figure 4.1 demonstrates the effect sizes for each TFT and COVID-19 severity-related meta-analysis.

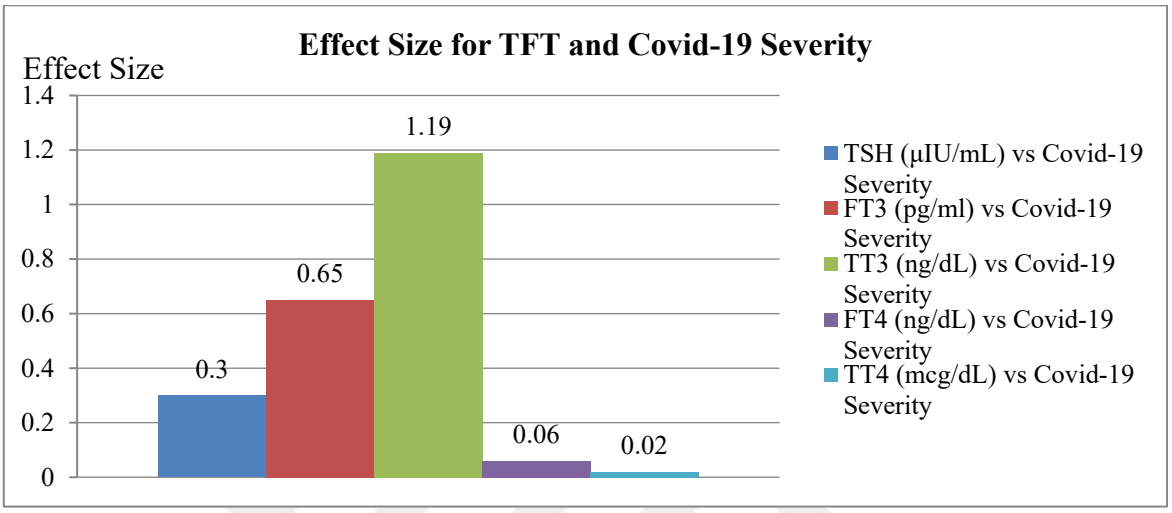

Figure 4.1. Effect size for TFT and COVID-19 severity

Second research question is trying to discover the comorbitity and COVID-19 severity relationship. Focused on diabeted and hypertension.

Comorbidity and COVID-19 Severity: Meta-Analysis 6: Comorbidity (Diabetes) vs COVID-19 Severity

Heterogeneity do not takes place ($Q = 3.16$, $p = 0.08$, $I^2 = 68\%$), fixed effect model is used. Effect size (ES) is measured by using odds ratio, ES = 0.46 (0.25 to 0.86). Among the mild/moderate and severe/critical groups odds ratio for diabetes is 0.46 greater in mild/moderate group within the 95% CI (0.25 to 0.86) which means the prognosis of COVID-19 is milder in patients having diabetes. Due to $p > 0.05$, there exists no statistically significant difference among the studies. According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. So, comorbidity (diabetes) have small effect on the COVID-19 severity. Publication bias is not observed regards to funnel plot and Begg test but observed regards to Egger test.

Comorbidity and COVID-19 Severity: Meta-Analysis 7: Comorbidity (Hypertension) vs COVID-19 Severity

Heterogeneity do not takes place (Q = 0.27, p = 0.6, I^2 = 0%), fixed effect model is used. Effect size (ES) is measured by using odds ratio, ES = 0.36 (0.2 to 0.68). Among the mild/moderate and severe/critical groups odds ratio for hypertension is 0.36 greater in mild/moderate group within the 95% CI (0.2 to 0.68) which means the prognosis of COVID-19 is milder in patients having hypertension. Due to p > 0.05, there exists no statistically significant difference among the studies. According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. So, comorbidity (hypertension) have small effect on the COVID-19 severity. Publication bias is not observed regards to funnel plot, Begg and Egger tests.

In conclusion, diabetes and hypertension comorbidities has small effect on the COVID-19 severity. Having diabetes has more effect on COVID-19 severity than having hypertension. Diabetes is very near the medium effect threshold (0.5 ≤ ES ≤ 0.8). Figure 4.2 demonstrates the effect sizes for each comorbidity and COVID-19 severity-related meta-analysis.

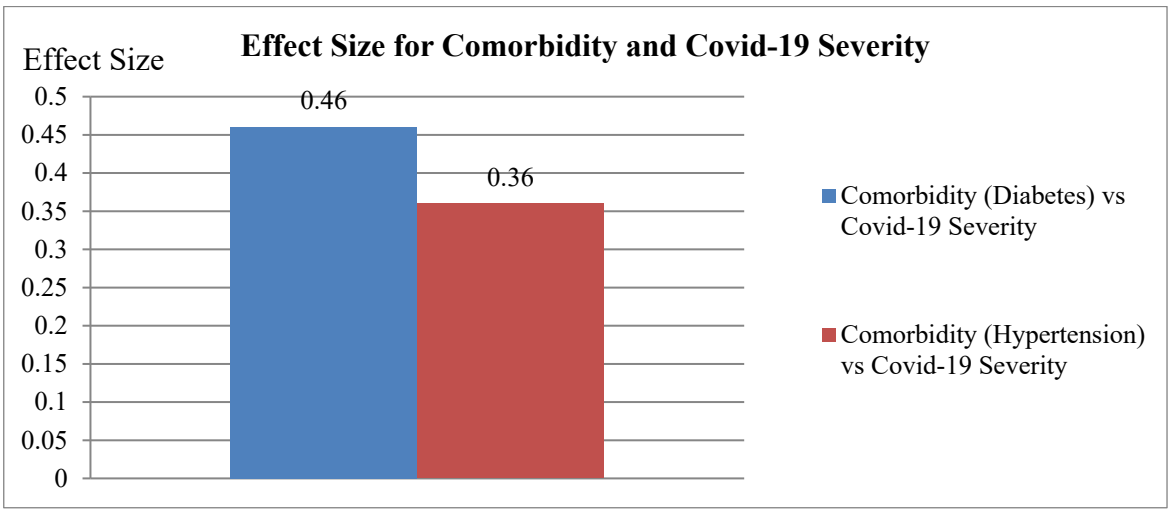

Figure 4.2. Effect size for comorbidity and COVID-19 severity

Third research question is trying to discover the TFT and mortality rates by COVID-19 relationship. Focused on TSH, FT3, TT3, FT4 hormone tests.

TFT and Mortality Rates by COVID-19: Meta-Analysis 8: TSH (μIU/mL) vs Mortality

Heterogeneity do not takes place (Q = 2.04, p = 0.73, I^2 = 0%), fixed effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 0.26 (0.05 to 0.47). Among the survived and deceased groups standardized mean TSH difference is 0.26 greater in the survived group within the 95% CI (0.05 to 0.47). According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. Due to p > 0.05, there exists no statistically significant difference among the studies. According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. So, TSH values have small effect on the COVID-19 mortality. Publication bias is not observed regards to funnel plot, Begg and Egger tests.

TFT and Mortality Rates by COVID-19: Meta-Analysis 9: FT3 (pg/mL) vs Mortality

Heterogeneity do not takes place (Q = 4.31, p = 0.77, I^2 = 54%) fixed effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 0.77 (0.49 to 1.06). Among the survived and deceased groups standardized mean FT3 difference is 0.77 greater in survived group within the 95% CI (0.49 to 1.06). According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,5 and 0,8 is evaluated as medium relative size. So, FT3 values have medium effect on the COVID-19 mortality. Publication bias is not observed regards to funnel plot, Begg and Egger tests.

TFT and Mortality Rates by COVID-19: Meta-Analysis 10: TT3 (ng/dL) vs Mortality

Heterogeneity do not takes place (Q = 1.16, p = 0.28, I^2 = 54%) fixed effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 0.73 (0.41 to 1.05). Due to p > 0.05, there exists no statistically significant difference among the studies. Among the survived and deceased groups standardized mean TT3 difference is 0.73 greater in survived group within the 95% CI (0.41 to 1.05). According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,5 and 0,8 is evaluated as medium relative size. So, TT3 values have medium effect on the COVID-19 mortality. Publication bias is not observed regards to funnel plot and Begg test, but observed in Egger test.

TFT and Mortality Rates by COVID-19: Meta-Analysis 11: FT4 (ng/dL) vs Mortality

Heterogeneity do not takes place (Q = 5.15, p = 0.27, I^2 = 22%) fixed effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 0.13 (-0.07 to 0.33). Due to p > 0.05, there exists no statistically significant difference among the studies. Among the survived and deceased groups standardized mean FT4 difference is 0.13 greater in survived group within the 95% CI (-0.07 to 0.33). Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. So, no relative size exists between the FT4 and COVID-19 mortality. Publication bias is not observed regards to funnel plot, Egger and Begg tests.

Fourth research question is trying to discover the thyroid dysfunction vs COVID-19 severity relationship. Thyroid dysfunction evaluated as non-thyroidal illness syndrome (NTIS), subclinical thyrotoxicosis, thyrotoxicosis, overt thyrotoxicosis, hypothyroidism, overt hypothyroidism, subclinical hypothyroidism, subnormal TSH, low FT3, high FT3, low FT4 and high FT4.

In conclusion, FT3, TT3 have medium; TSH has small; FT4 has very small or no effect on the COVID-19 severity. FT3 and TT3 are extremely near the large effect threshold (0.8 ≤ ES). Figure 4.3 demonstrates the effect sizes for each TFT and COVID-19 mortality-related meta-analysis.

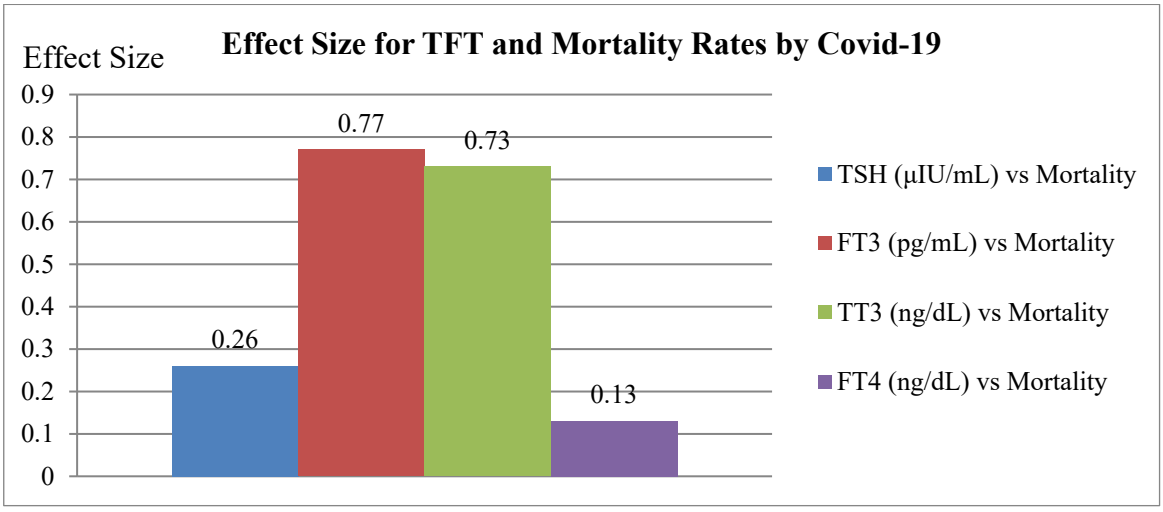

Figure 4.3. Effect size for TFT and COVID-19 mortality

Thyroid Dysfunction and COVID-19 Severity: Meta-Analysis 12: Thyroid Dysfunction vs COVID-19 Severity

Heterogeneity do not takes place (Q = 9.74, p = 0.14, I^2 = 38%), fixed effect model is used. Effect size (ES) is measured by using odds ratio, ES = 4.8 (3.03 to 7.6). Among the mild/moderate and severe/critical groups odds ratio for having thyroid dysfunction is 4.8 greater in severe/critical group within the 95% CI (3.03 to 7.6) which means the prognosis of COVID-19 is more severe in patients having thyroid dysfunction. Due to p > 0.05, there exists no statistically significant difference among the studies. According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is greater or equal than 0,8 is evaluated as a large relative size. So, having thyroid dysfunction have large effect on the COVID-19 severity. Publication bias is not observed regards to funnel plot, Begg and Egger tests.

Fifth research question is trying to discover the TFT and COVID-19 existence relationship. Focused on TSH, TT3, FT4 hormone tests. TSH, TT3, TT4 values for COVID-19 and non-COVID-19 patients. In addition, TSH value is investigated on COVID-19 and healthy individuals.

TFT and COVID-19 Existence: Meta-Analysis 13: TSH (μIU/mL) vs COVID-19 Existence

Heterogeneity do not takes place (Q = 2.73, p = 0.25, I^2 = 27%), fixed effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = -0.39 (-0.58 to -0.2). Among the COVID-19 and non-COVID-19 patients groups, standardized mean TSH difference is 0.39 greater in non-COVID-19 group within the 95% CI (-0.58 to -0.2). Due to p > 0.05, there exists no statistically significant difference among the studies. According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. So, TSH values have small effect on the COVID-19 existence. Publication bias is not observed regards to funnel plot, Begg and Egger tests.

TFT and COVID-19 Existence: Meta-Analysis 14: TT3 (ng/dL) vs COVID-19 Existence

Heterogeneity do not takes place (Q = 11.57, p = 0.003, I^2 = 83%), random effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = -

0.02 (-0.49 to 0.45). Among the COVID-19 and non-COVID-19 patients groups, standardized mean TT3 difference is 0.02 greater in non-COVID-19 group within the 95% CI (-0.49 to 0.45). Since, the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. So, no relative size exists between the TT3 values and COVID-19 existence. Publication bias is not observed regards to funnel plot, Begg and Egger tests.

TFT and COVID-19 Existence: Meta-Analysis 15: TT4 (mcg/dL) vs COVID-19 Existence

Heterogeneity do not takes place ($Q = 3.67$, $p = 0.06$, $I^2 = 73\%$) fixed effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = 0.47 (0.22 to 0.73). Due to $p > 0.05$, there exists no statistically significant difference among the studies. Among the COVID-19 and non-COVID-19 patients groups standardized mean TT4 difference is 0.47 greater in COVID-19 group within the 95% CI (0.22 to 0.73). According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is between 0,2 and 0,5 is evaluated as small relative size. So, TT4 values have small effect on the COVID-19 existence. Publication bias is not observed regards to funnel plot and Begg test, but observed in Egger test.

TFT and COVID-19 Existence: Meta-Analysis 16: TSH (μIU/mL) vs COVID-19/Healthy Status

Heterogeneity do not takes place ($Q = 2.53$, $p = 0.11$, $I^2 = 61\%$) fixed effect model is used. Effect size (ES) is measured by using standardized mean difference (SMD), ES = -1.08 (-1.3 to -0.87). Due to $p > 0.05$, there exists no statistically significant difference among the studies. Among the COVID-19 patients and non-COVID-19 healthy groups, standardized mean TSH difference is 1.08 greater in COVID-19 patients group within the 95% CI (-1.3 to -0.87). According to Cohen'(1988) "Rules-of-Thumb" for the standardized mean difference if the effect size is greater than 0.8, is evaluated as high relative size. So, TSH values have large effect on the COVID-19/healthy individuals. Publication bias is not observed regards to funnel plot and Begg test, but observed in Egger test.

In conclusion, TSH, TT4 have small and TT3 has very small or no effect on COVID-19 vs other diseased' patients. Despite that, TSH has large effect on COVID-19 patients vs

healthy individuals. Figure 4.4 demonstrates the effect sizes for each TFT and COVID-19 existence-related meta-analysis.

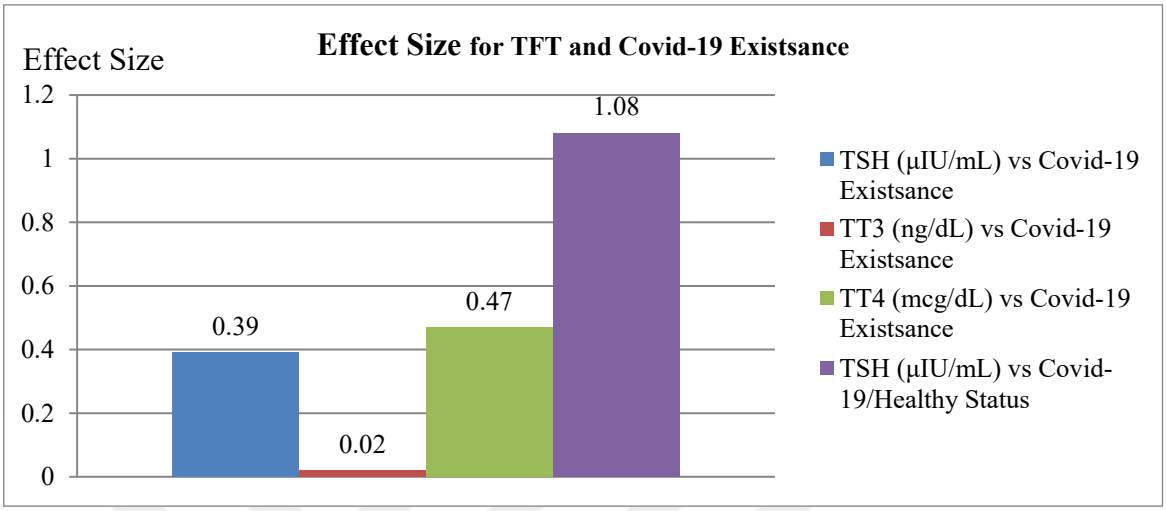

Figure 4.4. Effect size for TFT and COVID-19 existence

4.2. DISCUSSION & RELATED WORKS

While having the literature review also previous meta-analyses are examined coming from the used search queries in PubMed and Science Direct.

Llamas (2021) conducted a meta-analysis with 7 studies, 1183 patients from Turkey, Israel, China and Italy. Observed FT3 levels among the ICU and non-ICU admitted COVID-19 patients, FT3 levels between deceased/survived COVID-19 patients and low/normal FT3 levels on deceased/survived COVID-19 patients. He came up with a result as, both ICU-admitted (WMD = −0.91, 95% CI: −1.08; −0.73) and deceased (WMD = −1.33, 95% CI: −2.03; −0.64) COVID-19 patients had lower FT3 levels than non-ICU admitted and survived ones. So, FT3 levels are dramatically lower in severe COVID-19 patients than in non-severe. [70]

Chen (2022) conducted a meta-analysis with 20 studies, 3609 patients from Italy, China, Turkey, UK, Israel, Greece, India, Brasil, Korea and Spain. Observed relationship between the FT3, TSH, FT4 levels vs COVID-19 severity and mortality. He came up with a result as, thyroid disorders are more common in severe COVID-19 patients and mortality rates are higher; FT3, TSH levels are lower in severe COVID-19 patients; FT4 levels do not

have significant effect on COVID-19 severity. Also, deceased/survived status shows the same characteristics. In conclusion, low FT3 and TSH levels may increase the mortality rates and causes more severe COVID-19. [71]

Permana (2022) conducted a meta-analysis and meta-regression with 20 studies, 24734 patients. Observed relationship between the comorbitities (coronary artery disease/cardiovascular disease (CAD/CVD), chronic obstructive pulmonary disease (COPD), diabetes mellitus (DM), hypertension) vs severity, hospitalization, mortality and ICU admission of COVID-19. He came up with the result as, thyroid disorders increase the COVID-19 severity, mortality, ICU admission and hospitalization. Meta-regression results showed age, hypertension but not DM, CAD/CVD, gender, COPD, obesity affects the association. [72]

Damara (2021) conducted a meta-analysis and meta-regression with 21 studies, 31339 patients. Observed relationship between thyroid disorders vs severity, hospitalization, mortality and ICU admission of COVID-19. He came up with the result as, thyroid disorders increase the COVID-19 severity, mortality, ICU admission and hospitalization. Meta-regression results showed age affects the association. Hypothyroidism and not normal thyroid –not hyperthyroidism- are associated with poor COVID-19 outcomes. [73]

All in all, Llamas (2021) work has resulted in ICU-admitted, deceased and severe COVID-19 patients having lower FT3 levels. Present study indicates that FT3 has medium effect on COVID-19 severity and medium-large effect on COVID-19 related-mortality. Among the mild/moderate and severe/critical groups standardized mean FT3 difference is 0.65 greater in mild/moderate group within the 95% CI (0.31 to 0.99). Among the survived and deceased groups standardized mean FT3 difference is 0.77 greater in survived group within the 95% CI (0.49 to 1.06). Present study' results are similar to Llamas' (2021) result.

Chen (2022) work has resulted in thyroid disorders increasing the mortality and severity of COVID-19, low FT3 and TSH levels may increase the mortality rates and causes more severe COVID-19 –FT4 has no effect-. Present study is just able to assess thyroid disorder and COVID-19 severity due to the limitation in data. Indicates that thyroid dysfunction have large effect on the COVID-19 severity. Among the mild/moderate and severe/critical groups odds ratio for having thyroid dysfunction 4.8 greater in severe/critical group within the 95% CI (3.03 to 7.6) which means the prognosis of COVID-19 is more severe in

patients having thyroid dysfunction. TSH values have no effect on the COVID-19 severity. Among the mild/moderate and severe/critical groups standardized mean TSH difference is 0.3 greater in mild/moderate group within the 95% CI (-0.05 to 0.65). Since, in the forest plot the CI's of studies passes through the zero axis, interpreting the effect size is not meaningful. TSH values have small effect on the COVID-19 related mortality. Among the survived and deceased groups standardized mean TSH difference is 0.26 greater in the survived group within the 95% CI (0.05 to 0.47). Similar to Chen' (2022) output, the current study presents FT4 have no effect on COVID-19 severity and mortality but, FT3 related outcomes similar to Chen' (2022) study. TSH related outcomes are contradict with Chen' (2022).

Permana (2022) and Damara (2021) studies have resulted in the same conclusion: thyroid disorders increase the COVID-19 severity, mortality, ICU admission and hospitalization. The present study is just able to assess thyroid disorder and COVID-19 severity due to the limitation in data and ended with similar outputs to them. Indicates that thyroid dysfunction have large effect on the COVID-19 severity. Among the mild/moderate and severe/critical groups odds ratio for having thyroid dysfunction 4.8 greater in severe/critical group within the 95% CI (3.03 to 7.6) which means the prognosis of COVID-19 is more severe in patients having thyroid dysfunction.

In conclusion, it is shown that thyroid dysfunction patients tend to undergo more severe COVID-19. TSH, FT4, TT4 values have no effect on COVID-19 severity. Also, FT4 values was not affect COVID-19-related mortality. Patients with higher TF3 and TT3 levels tend to undergo milder COVID-19. Moreover, the chance of surviving from COVID-19 gets increases.

4.3. LIMITATIONS

There were several limitations of the presented meta-analyses work. First of all, in publication bias assessment, fewer studies containing meta-analyses, evaluation of either having bias or not, changes for each funnel plot, Egger and Begg tests. According to Nik (2012), in moderate bias both Egger and Begg acts well but their precision is low in small meta-analysis. [74]

Second of all, quality score of all meta-analyses was resulted as “low”. Including observational studies (cohort, cross-sectional) and due to the limitations in terms of time, some studies have been done with insufficient or unbalanced sample groups, populations, demographic structures etc. resulted as lower reliability levels.

Third of all, since each study focused on different aim and outcome, trying to combine them is difficult and resulted as small sample sizes are low rather than previous meta-analyses works. Also, thyroid hormone units, type of statistical measure (mean, median) given and COVID-19 severity classification (mild, moderate, severe etc.) differs from study to study. Converting them to a single measure is the biggest challenge during this work.

REFERENCES

1. Field, Andy P., and Raphael Gillett. How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology* 63.3 (2010): 665-694.
2. Aslanyürek Sezer, Emine. *Ortaokul Öğrencilerinin Eleştirel Düşünme Becerisi Üzerine Bir Meta-Analiz Çalışması*. Master Thesis, Pamukkale Üniversitesi, 2021.
3. Bulut, Harun. *Klinik Çalışmalarda Meta Analizi Ve Meta Regresyon: Bir Uygulama*. Master Thesis, Biruni Üviversitesi, 2021.
4. Kaya, Serap. *Öğrencilerin Okuma Alışkanlıklarını Etkileyen Faktörler Üzerine Bir Meta-Analiz Çalışması*. Master Thesis, Kırşehir Ahi Evran Üniversitesi, 2021.
5. Çarkungöz, E. , Ediz, B. Meta Analizi . *Uludağ Üniversitesi Veteriner Fakültesi Dergisi* 28 (2009): 33-37 https://dergipark.org.tr/tr/pub/uluvfd/issue/13524/163600
6. Abramson JH. Making Sense of Data. 2nd edition. New York: Oxford University Press; 1994.
7. Glass, Gene V. Meta-analysis at middle age: a personal history. *Research synthesis methods* vol. 6,3 (2015): 221-31. doi:10.1002/jrsm.1133
8. Bailar, J. C., & Hoaglin, D. C. (2012). Medical uses of statistics. John Wiley & Sons. Bailey, K. R. (1987). Inter-study differences: How should they influence the interpretation and analysis of results? Statistics in Medicine, 6(3), 351–358.
9. Şen, Sedat. Meta-Analiz. 1 Nov. 2016. PowerPoint Slides.
10. West SL, Gartlehner G, Mansfield AJ, et al. Comparative Effectiveness Review Methods: Clinical Heterogeneity [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2010 Sep. Table 7, Summary of common statistical approaches to test for heterogeneity. Available from: https://www.ncbi.nlm.nih.gov/books/NBK53317/table/ch3.t2/8
11. Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). Introduction to meta-analysis. United Kingdom: Wiley
12. Huedo-Medina, Tania B et al. Assessing heterogeneity in meta-analysis: Q statistic or I2 index?. *Psychological methods* vol. 11,2 (2006): 193-206. doi:10.1037/1082-989X.11.2.193
13. Hakverdi, G. Meta Analizi ve Genetik Çalışmalarda Bir Uygulama. *İnönü University*, 2017.

14. Spineli, Loukia M., and Nikolaos Pandis. Meta-analysis: fixed-effect model. *American journal of orthodontics and dentofacial orthopedics* 157.1 (2020): 134-137.
15. Statistics Solutions. (2013). Effect Size. Available from https://www.statisticssolutions.com/academic-solutions/resources/directory-of-statistical-analyses/effect-size/
16. Özgül, Burcu. *Probleme Dayalı Öğrenme Yaklaşımının Etkililiği: Bir Meta-Analiz Çalışması*. Master Thesis, Burdur Mehmet Akif Ersoy Üniversitesi, 2021.
17. Tak, Ayşegül Yabacı. Etki büyüklüğü yöntemlerinin karşılaştırılması. Diss. *Bursa Uludag University* (Turkey), 2021.
18. Cohen, J. (1988). Statistical power analysis for the behavioral sciences. US: Lawrence Erlbaum Associates
19. Odds Ratio. *Psych*, Available from psychscenehub.com/psychpedia/odds-ratio-2/.
20. Frost, Jim. Relative Risk: Definition, Formula & Interpretation. *Statistics By Jim*, statisticsbyjim.com/probability/relative-risk/.
21. (2008). Pearson's Correlation Coefficient. In: Kirch, W. (eds) Encyclopedia of Public Health. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-5614-7_2569
22. Calkins, Keith G. Correlation Coefficients. 18 Jul. 2005. Applied Statistics - Lesson 5.
23. Çarkungöz, Ender. *Meta Analizinin Veteriner Hekimlikte Uygulanması*. Master Thesis, Uludağ Üniversitesi, 2010.
24. Olkin I. Meta- Analysis: Current Issues In Research Synthesis. Statistics in Medicine, 15: 1253-1257, 1996.
25. Jenicek M. Meta Analysis in Medicine: Where we are and where we want to go. *Journal of Clinical Epidemiolgy,* 42(1): 35-44, 1989.
26. Mcelvenny D, Armstrong Bg, Järup L, Higgins Jpt. Meta Analysis in Occupational Epidemiology: A Review of Practice. *Occupational Medicine*, 54: 336-344, 2004.
27. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994, 50: 1088-
28. Lewis, S, and M Clarke. Forest plots: trying to see the wood and the trees. BMJ (Clinical research ed.) vol. 322,7300 (2001): 1479-80. doi:10.1136/bmj.322.7300.1479
29. Harrer, Mathias, et al. What Is a Forest Plot? *Bookdown*, bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/forest.html#what-is-a-forest-plot.

30. Basics of COVID-19. *Centers for Disease Control and Preventation*, 4 Nov. 2021, www.cdc.gov/coronavirus/2019-ncov/your-health/about-COVID-19/basics-COVID-19.html.
31. Britannica, The Editors of Encyclopaedia. thyroid gland. *Encyclopedia Britannica*, 2 Sep. 2022, https://www.britannica.com/science/thyroid-gland.
32. Tiroit Bezi ve Fonksiyonları. *Ceylan Hastanesi*, Available from www.ceylanhastanesi.com/tiroid-bezi-ve-fonksiyonlari/.
33. Tiroid Hastalıkları ve Tiroid Nodülleri. *Acıbadem*, 4 Oct. 2019, Available from www.acibadem.com.tr/ilgi-alani/tiroid-hastaliklari-ve-tiroid-nodulleri/#tani-yontemleri.
34. Thyrotoxicosis. *Cleveland Clinic*, 23 Aug. 2021, Available from my.clevelandclinic.org/health/diseases/21741-thyrotoxicosis.
35. Charma, Anu, and Stan, Marius N. Thyrotoxicosis: Diagnosis and Management. *Mayo Clinic*, 2018, https://doi.org/10.1016/j.mayocp.2018.10.011.
36. Gessl, Alois et al. Thyroid disorders. *Handbook of experimental pharmacolog,* 214 (2012): 361-86. doi:10.1007/978-3-642-30726-3_17
37. Hyperthyroidism (overactive thyroid). *Mayo Clinic*, 14 Nov. 2020, Available from www.mayoclinic.org/diseases-conditions/hyperthyroidism/symptoms-causes/syc-20373659.
38. Taşkale, Mustafa Gürkan. Tiroid nedir? Tiroid belirtileri nelerdir?"*Medical Park*, 22 Jun. 2020, Available from www.medicalpark.com.tr/tiroid/hg-2285.
39. Thyroid Disease. *Cleveland Clinic*, 19 Apr. 2020, Available from my.clevelandclinic.org/health/diseases/8541-thyroid-disease.
40. Braverman, Lewis E, Kopp Peter, David S. Cooper, Sidney C. Werner, and Sidney H. Ingbar. *Werner & Ingbar's the Thyroid: A Fundamental and Clinical Text.* Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins, 2021
41. Thyroiditis. *American Thyroid Association*, Available from www.thyroid.org/thyroiditis/#:~:text=If%20the%20thyroiditis%20causes%20rapid,hyperthyroidism%20(see%20Hyperthyroidism%20brochure).
42. Hypothyroidism (underactive thyroid). *Mayo Clinic*, 19 Nov. 2020, Available from www.mayoclinic.org/diseases-conditions/hypothyroidism/symptoms-causes/syc-20350284.

43. Budak, F. , Korkmaz, Ş. COVID-19 Pandemi Sürecine Yönelik Genel Bir Değerlendirme: Türkiye Örneği. *Sosyal Araştırmalar ve Yönetim Dergisi (2020)*: 62-79 https://dergipark.org.tr/en/pub/sayod/issue/54364/738657
44. Coronavirus, COVID-19. *Cleveland Clinic*, 1 Mar. 2022, Available from my.clevelandclinic.org/health/diseases/21214-coronavirus-COVID-19.
45. WHO Coronavirus (COVID-19) Dashboard. *World Health Organizaiton*, Available from COVID19.who.int/table.
46. COVID-19 Nedir ? *T.C. Sağlık Bakanlığı COVID-19 Bilgilendirme Platformu*, Available from COVID19.saglik.gov.tr/TR-66300/COVID-19-nedir-.html.
47. Coronavirus disease 2019 (COVID-19). *Mayo Clinic*, 29 Oct. 2022, Available from www.mayoclinic.org/diseases-conditions/coronavirus/diagnosis-treatment/drc-20479976.
48. Hariyanto, Timotius Ivan, and Andree Kurniawan. Thyroid disease is associated with severe coronavirus disease 2019 (COVID-19) infection. *Diabetes & metabolic syndrome* vol. 14,5 (2020): 1429-1430. doi:10.1016/j.dsx.2020.07.044
49. Trimboli, P., Camponovo, C., Scappaticcio, L. et al. Thyroid sequelae of COVID-19: a systematic review of reviews. *Rev Endocr Metab Disord* 22, 485–491 (2021). https://doi.org/10.1007/s11154-021-09653-1
50. Akgöz, S., Ercan, İ. ve Kan, İ. (2004). Meta-analizi. *Uludağ Üniversitesi Tıp Fakültesi Dergisi,* 30(2), 107-112
51. PubMed.gov. *National library of Medicine*, Available from pubmed.ncbi.nlm.nih.gov/.
52. Science Direct. *Science Direct*, Available from www.sciencedirect.com/.
53. Ahn, Jiyeon et al. Thyroid Hormone Profile and Its Prognostic Impact on the Coronavirus Disease 2019 in Korean Patients. *Endocrinology and metabolism (Seoul, Korea)* vol. 36,4 (2021): 769-777. doi:10.3803/EnM.2021.1109
54. Gao, W et al. Thyroid hormone concentrations in severely or critically ill patients with COVID-19. *Journal of endocrinological investigation* vol. 44,5 (2021): 1031-1040. doi:10.1007/s40618-020-01460-w
55. Vassiliadi, Dimitra Argyro et al. Thyroid hormone alterations in critically and non-critically ill patients with SARS-CoV-2 infection. *Endocrine connections* vol. 10,6 646-655. 17 Jun. 2021, doi:10.1530/EC-21-0029

56. Malik, Jahanzeb et al. Thyroid function analysis in COVID-19: A retrospective study from a single center. *PloS one* vol. 16,3 e0249421. 30 Mar. 2021, doi:10.1371/journal.pone.0249421
57. Lui, David Tak Wai et al. Thyroid Dysfunction in Relation to Immune Profile, Disease Status, and Outcome in 191 Patients with COVID-19. *The Journal of clinical endocrinology and metabolism* vol. 106,2 (2021): e926-e935. doi:10.1210/clinem/dgaa813
58. Baldelli, R et al. Thyroid dysfunction in COVID-19 patients. *Journal of endocrinological investigation* vol. 44,12 (2021): 2735-2739. doi:10.1007/s40618-021-01599-0
59. Bonakdaran, Shokoufeh et al. The Prognostic Role of Metabolic and Endocrine Parameters for the Clinical Severity of COVID-19. *Disease markers* vol. 2022 5106342. 20 Jan. 2022, doi:10.1155/2022/5106342
60. Güven, Mehmet, and Hamza Gültekin. The prognostic impact of thyroid disorders on the clinical severity of COVID-19: Results of single-centre pandemic hospital. *International journal of clinical practice* vol. 75,6 (2021): e14129. doi:10.1111/ijcp.14129
61. Dincer Yazan, Ceyda et al. The Association of Thyroid Hormone Changes with Inflammatory Status and Prognosis in COVID-19. *International journal of endocrinology* vol. 2021 2395212. 13 Aug. 2021, doi:10.1155/2021/2395212
62. Lui, David Tak Wai et al. Role of non-thyroidal illness syndrome in predicting adverse outcomes in COVID-19 patients predominantly of mild-to-moderate severity. *Clinical endocrinology* vol. 95,3 (2021): 469-477. doi:10.1111/cen.14476
63. Sen, Kaushik et al. Thyroid Function Test in COVID-19 Patients: A Cross-Sectional Study in a Tertiary Care Hospital. *Indian journal of endocrinology and metabolism* vol. 24,6 (2020): 532-536. doi:10.4103/ijem.IJEM_779_20
64. Nakamura, Sachiko et al. Analysis of thyroid function in Japanese patients with coronavirus disease 2019. *Endocrine journal* vol. 69,6 (2022): 643-648. doi:10.1507/endocrj.EJ21-0609
65. Wang, Weibin et al. Thyroid Function Abnormalities in COVID-19 Patients. *Frontiers in endocrinology* vol. 11 623792. 19 Feb. 2021, doi:10.3389/fendo.2020.623792

66. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. Available from www.math.hkbu.edu.hk/~tongt/papers/median2mean.html.
67. Weighted Average Calculator. *RapidTables*, Available from www.rapidtables.com/calc/math/weighted-average-calculator.html.
68. GRADEpro GDT*GRADEpro GDT*, Available from gdt.gradepro.org/.
69. Cochrane Community. *GRADEpro GDT*, Available from community.cochrane.org/help/tools-and-software/gradepro-gdt.
70. Llamas, Michael et al. Low free-T3 serum levels and prognosis of COVID-19: systematic review and meta-analysis. *Clinical chemistry and laboratory medicine* vol. 59,12 1906-1913. 11 Aug. 2021, doi:10.1515/cclm-2021-0805
71. Chen, Yiru et al. The Association Between COVID-19 and Thyroxine Levels: A Meta-Analysis. *Frontiers in endocrinology* vol. 12 779692. 4 Jan. 2022, doi:10.3389/fendo.2021.779692
72. Permana, Hikmat et al. The prognostic values of thyroid disorders in predicting COVID-19 composite poor outcomes: A systematic review and meta-analysis. *Diabetes & metabolic syndrome* vol. 16,5 (2022): 102464. doi:10.1016/j.dsx.2022.102464
73. Damara, Fachreza Aryo et al. Thyroid disease and hypothyroidism are associated with poor COVID-19 outcomes: A systematic review, meta-analysis, and meta-regression. *Diabetes & metabolic syndrome* vol. 15,6 (2021): 102312. doi:10.1016/j.dsx.2021.102312
74. Nik Ruzni Nik Idris, 2012. A Comparison of Methods to Detect Publication Bias for Meta-analysis of Continuous Data. *Journal of Applied Sciences, 12: 1413-1417.*

APPENDIX A: QUALITY SCORE ASSESSMENT OF THE STUDIES

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	[intervention]	[comparison]	Relative (95% CI)	Absolute (95% CI)		
TSH vs Covid-19 Severity												
9	observational studies	not serious	not serious	not serious	not serious	none	535	566	-	SMD **0.3 higher** (0.05 lower to 0.65 higher)	⨁⨁◯◯ Low	
TF3 vs Covid-19 Severity												
5	observational studies	not serious	not serious	not serious	not serious	none	433	324	-	SMD **0.65 higher** (0.31 higher to 0.99 higher)	⨁⨁◯◯ Low	
TT3 vs Covid-19 Severity												
5	observational studies	not serious	not serious	not serious	not serious	none	139	260	-	SMD **0.96 higher** (0.72 higher to 1.2 higher)	⨁⨁◯◯ Low	
FT4 vs Covid-19 Severity												
7	observational studies	not serious	not serious	not serious	not serious	none	477	453	-	SMD **0.07 higher** (0.07 lower to 0.22 higher)	⨁⨁◯◯ Low	
TT4 vs Covid-19 Severity												
2	observational studies	not serious	not serious	not serious	not serious	none	64	44	-	SMD **0.02 higher** (0.38 lower to 0.42 higher)	⨁⨁◯◯ Low	
Diabetes vs Covid-19 Severity												
2	observational studies	not serious	not serious	not serious	not serious	none	18/91 (19.8%)	44/126 (34.9%)	**OR 0.54** (0.16 to 1.79)	**125 fewer per 1.000** (from 270 fewer to 141 more)	⨁⨁◯◯ Low	
Hypertension vs Covid-19 Severity												
2	observational studies	not serious	not serious	not serious	not serious	none	28/91 (30.8%)	64/126 (50.8%)	**OR 0.36** (0.19 to 0.67)	**237 fewer per 1.000** (from 344 fewer to 99 fewer)	⨁⨁◯◯ Low	
TSH vs Mortality												
5	observational studies	not serious	not serious	not serious	not serious	none	459	126	-	SMD **0.26 higher** (0.05 higher to 0.47 higher)	⨁⨁◯◯ Low	
FT3 vs Mortality												
3	observational studies	not serious	not serious	not serious	not serious	none	323	68	-	SMD **0.77 higher** (0.49 higher to 1.06 higher)	⨁⨁◯◯ Low	
TT3 vs Mortality												
2	observational studies	not serious	not serious	not serious	not serious	none	131	58	-	SMD **0.72 higher** (0.38 higher to 1.07 higher)	⨁⨁◯◯ Low	

Thyroid Dysfunction vs Covid-19 Severity

7	observational studies	not serious	not serious	not serious	not serious	none	109/711 (15.3%)	228/359 (63.5%)	**OR 4.80** (3.03 to 7.60)	**258 more per 1.000** (from 206 more to 295 more)	⨁⨁◯◯ Low	
TSH vs Covid-19 Existance												
3	observational studies	not serious	not serious	not serious	not serious	none	234	213	-	SMD **0.38 lower** (0.6 lower to 0.15 lower)	⨁⨁◯◯ Low	
TT3 vs Covid-19 Existance												
3	observational studies	not serious	not serious	not serious	not serious	none	234	213	-	SMD **0.1 higher** (0.09 lower to 0.29 higher)	⨁⨁◯◯ Low	
TT4 vs Covid-19 Existance												
2	observational studies	not serious	not serious	not serious	not serious	none	132	56	-	SMD **0.35 higher** (0.14 lower to 0.83 higher)	⨁⨁◯◯ Low	
TSH vs Covid-19/Healty Individuals												
2	observational studies	not serious	not serious	not serious	not serious	none	130	827	-	SMD **1.2 lower** (1.68 lower to 0.72 lower)	⨁⨁◯◯ Low	

CI: confidence interval; **OR:** odds ratio; **SMD:** standardised mean difference