Understanding the Causes of Stress: Induced Cardiomyopathy Using Bioinformatics

ABSTRACT

Background

Stress-induced cardiomyopathy is a heart disease that is caused by stress. The main target group for this disease is post-menopausal women. It causes the left ventricle to balloon, which mimics symptoms of a heart attack including chest pain, shortness of breath, and fatigue. The problem with this disease is that we don't know much about the causes on a genetic level. Therefore, the goal of this research is to find genes that are highly expressed in Stress-induced cardiomyopathy.

Methods

In this study, **GSE95368** human sample dataset from the NCBI GEO2R bioinformatics database was collected and analyzed with GEO2R to identify differentially expressed genes (DEGs). The samples were grouped into three; those infected with stress-induced cardiomyopathy, with Acute myocardial infarction, and from healthy human samples. The top 30 DEGs were then analyzed using GO and KEGG pathway enrichment analysis from SR Plot bioinformatics tool to determine their potential functions.

Results

We identified 30 top DEGs based on p-value and fold change split into 15 upregulated and 15 downregulated genes. The GO analysis showed high enrichment in the biological processes ontology with maintenance of location and plasma lipoprotein particle organization being prominent functions. The GO showed high enrichment in molecular functions with heparin and glycosaminoglycan binding as the highest functions. The KEGG results showed that cholesterol metabolism, coagulation, and complement cascade pathways were the most important pathways. Common genes found in both results were the APOE, F2, and APOB genes

Conclusion

The study revealed that the top three genes related to Stress-Induced Cardiomyopathy are APOE, F2, and APOB. Further observations of these genes in a laboratory setting can be used to determine potential early diagnosis of Stress-Induced Cardiomyopathy.

Keywords: NCBI GEO2R, Stress-induced cardiomyopathy, SR Plot, KEGG, Gene Ontology, APOE, F2, and APOB

INTRODUCTION

Stress-induced cardiomyopathy is a heart disease that is caused by stress. It happens when someone is stressed and their left ventricle begins to weaken, which makes it harder for the heart to pump blood (1). The effects are very similar to a heart attack including chest pain, shortness of breath, and fatigue (2). However, unlike a heart attack instead of clogged arteries, the left ventricle will balloon out. This disease is important to learn about because it is potentially life-threatening and unknown due to its rarity.

The main problem of Stress-Induced Cardiomyopathy is understanding the causes of how it can affect certain people (1). We want to understand better what is specifically responsible for causing the disease to know who is at higher risk of getting it. The scientific question being investigated is ‘What types of genes are responsible for causing Stress-Induced Cardiomyopathy? For example, are these genes involved in stress?’

So far it is known that Stress-Induced Cardiomyopathy is an emotionally triggered heart disease. It is also known that family history and catecholamine (stress hormone) signals can play a factor in the disease (3). The main demographic for this disease is postmenopausal women (4). Most cases of Stress-Induced Cardiomyopathy are typically benign, but there have been cases of death. However, a patient with Stress-Induced Cardiomyopathy will most likely return to their normal self within about a few weeks or months with the help of beta-blockers to reduce stress hormones (5).

In bioinformatics, it is known that genes that are responsible for stress are involved in causing this disease (5). The challenge with Stress-Induced Cardiomyopathy is finding out the specific genes that are responsible for causing the disease, or how they can be triggered. Also, there is not a lot of bioinformatics research on the disease due to its rarity and unknown causes.

The goal of this research is to understand which genes are important in causing Stress-Induced Cardiomyopathy and how they are activated. I hypothesize that the genes that are important in causing Stress-Induced Cardiomyopathy are commonly activated when someone feels stress.

This research is important because it can help other people understand what Stress-Induced Cardiomyopathy is, and how it can be prevented or diagnosed before any damage is to occur.

METHODS

Data Collection and Analysis of GEO2R Data

NCBI is the National Center for Biotechnology Information that contains a great amount of biological information (6). A research paper, found on NCBI, focused on Stress-Induced Cardiomyopathy and its comparisons to healthy patients and patients with Acute Myocardial Infarction. This research paper included genes that were able to be analyzed by GEO2R. GEO2R is a tool within NCBI that is used to help visualize genetic information (6). The dataset collected was **GSE95368** and it was analyzed in GEO2R by first defining these three groups: Stress-Induced Cardiomyopathy, Acute Myocardial Infarction, and Healthy (Figure 1).

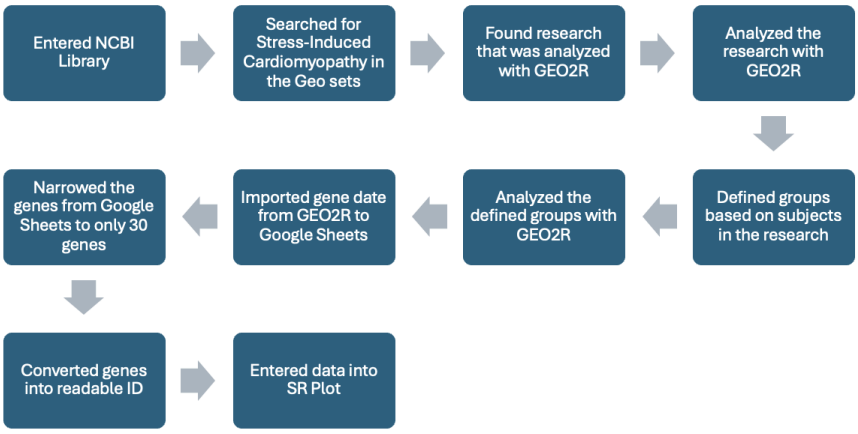

Figure 1. Research Methodology: The steps and bioinformatics used in this study.

Identification of the Top Differentially Expressed Genes

From the GEO2R results, 1225 genes were observed to be differentially expressed. To identify the 30 most significant differentially expressed genes, statistical analysis was applied. This process identified different pValues and Fold changes to prioritize the most important genes based on their differential expression across samples. The pValues were narrowed down by finding genes that had a pValue of less than 0.05. The fold changes were narrowed by finding the top 15 highest fold changes and the bottom 15 lowest fold changes.

Data Analysis Using SRPlot, KEGG, and GO Bioinformatics Tools

Then SRPlot, KEGG, and GO bioinformatics tools and databases were utilized to analyze the potential functions of these top genes. SR Plot is a bioinformatics tool that is used to create graphs to analyze gene pathways (7). SR Plot divides these graphs into two different results, KEGG and GO. KEGG is a bioinformatics tool that collects and organizes information on genes and their processes (8). GO is another bioinformatics tool that organizes genes by their functions (9). Altogether, these bioinformatics tools helped uncover the potential roles of the genes in their functions and the different pathways that they might be involved in.

RESULTS

Identification of Differentially Expressed Genes

The first bioinformatics tool that was used to identify differentially expressed genes (DEGs)was from GEO2R on NCBI. From the first two results, some genes were expressed differently in healthy patients and those with Stress-Induced Cardiomyopathy (Figures 1 and 2).

In the volcano plot ((Figure 2a) the red dots represent upregulated genes, meaning that they are much more expressed. The blue dots represent downregulated genes, meaning that they are down-expressed. The black dots represent genes that represent the same amount of expression in a healthy person and one with Stress-Induced Cardiomyopathy (Figure 2a).

In our study, the total number of genes that were differentially expressed was 1225 different genes (Figure 2b). The Venn diagram shows that between healthy, acute stress-induced cardiomyopathy and acute myocardial infarction, both diseases do not share many of the same genes, but there is an overlap of 44 between the Acute Stress-Induced Cardiomyopathy vs Healthy group and the Acute Myocardial Infarction vs Healthy group (Figure 2b).

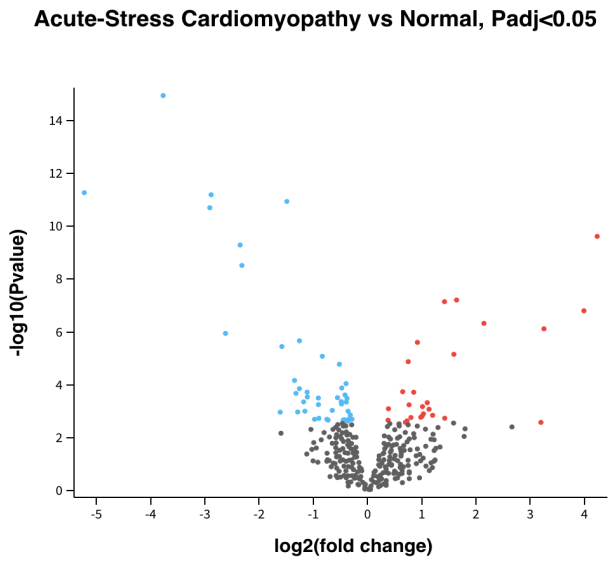

Figure 2a. The volcano plot shows how the genes are being expressed differently between Stress-Induced Cardiomyopathy and healthy individuals. The red dots represent upregulated genes, the blue dots represent down-regulated genes, and the black dots represent an equal amount of expression.

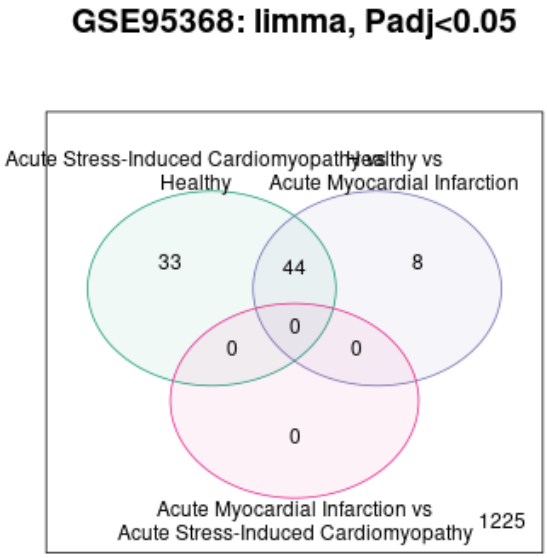

Figure 2b. The Venn diagram shows the genes that are shared between Stress-Induced Cardiomyopathy, Acute Myocardial Infarction, and healthy patients. Out of 1225 genes that are expressed there is an overlap of genes within the Stress-Induced Cardiomyopathy vs Healthy group and the Healthy vs Acute Myocardial Infarction group.

Identification of Top 30 Statistically Significant Differentially Expressed Genes (DEGs)
The genes that were downloaded from the study were narrowed down based on the p-value and fold change. First, they were selected by finding genes that had a p-value lower than 0.05. Then they were narrowed further by finding the lowest and highest fold changes Genes with a log2 fold change (|FC|) greater than 1 were classified as differentially expressed. In the end, the genes were narrowed to only 30. These were based on the top 15 highest logFC and the bottom 15 lowest logFC. There were 15 upregulated genes and 15 downregulated genes after the statistical analysis (Top 30 Genes).

Potential Functions and Enrichment of the Identified Genes and Pathways
In the KEGG results from SR Plot, the complement and coagulation cascade pathway, and cholesterol metabolism were both identified as significant in the results (Figure 3b).. The complement and coagulation cascade pathway and cholesterol metabolism had the highest enrichment scores. From the KEGG results, one gene that stood out were the ApoE and

ApoB-100 genes that were found in the complement and coagulation cascades pathway (Figure 3a). Some other genes that stood out were Serpin, F2, PAI, PCI, and C3 which were found in the cholesterol metabolism pathway (Figure 3b).

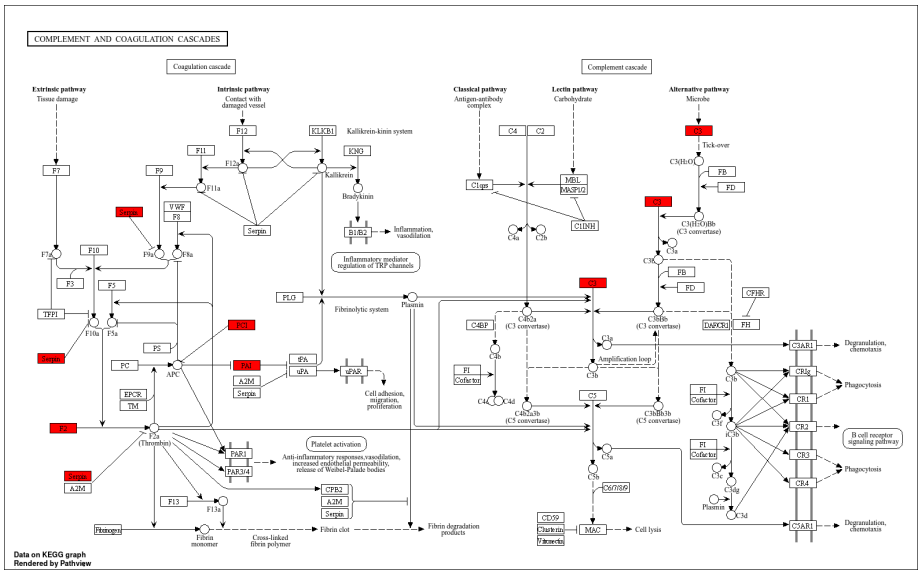

Figure 3a. This image represents the complement and coagulation cascade pathway and the genes significantly involved in the KEGG result in red. Significant genes include Serpin, C3, PAI, and F2.

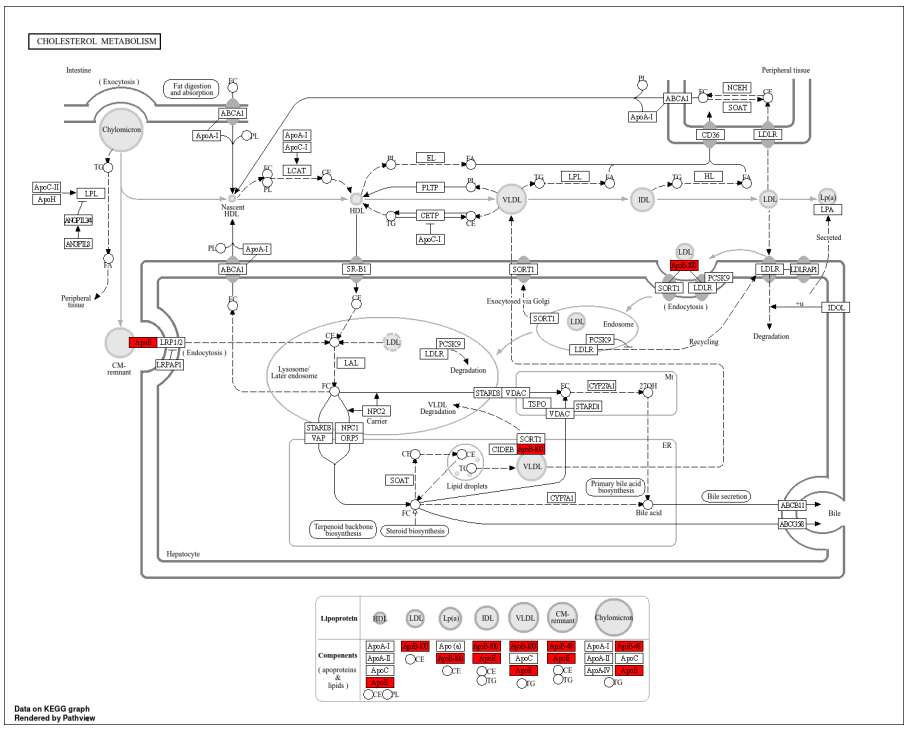

Figure 3b. This image represents the Cholesterol metabolism pathway and the genes significantly involved in the KEGG result in red. Significant genes include ApoB-100, ApoE, and ApoB-48

The Gene Ontology (GO) is separated into three ontologies: biological processes, molecular functions, and cellular components. The results showed that all of the ontologies were relatively equal in enrichment. However, some pathways with a higher enrichment score were maintenance of location and plasma lipoprotein particle organization, which is found in the biological processes ontology; as well as heparin and glycosaminoglycan binding, found in molecular functions (Figure 4a).

From the GO results, some genes that stood out in the multicellular function were the SFRP1, APOE, SERPINA5, SAA1, F2, and APOB genes. Some genes that stood out in the biological processes were the APOE, CRP, ALB, PRKACA, F2, APOB, and C3 genes (Figure 4b). Common genes that stood out in both MP and BP ontology were the APOE, F2, and APOB genes.

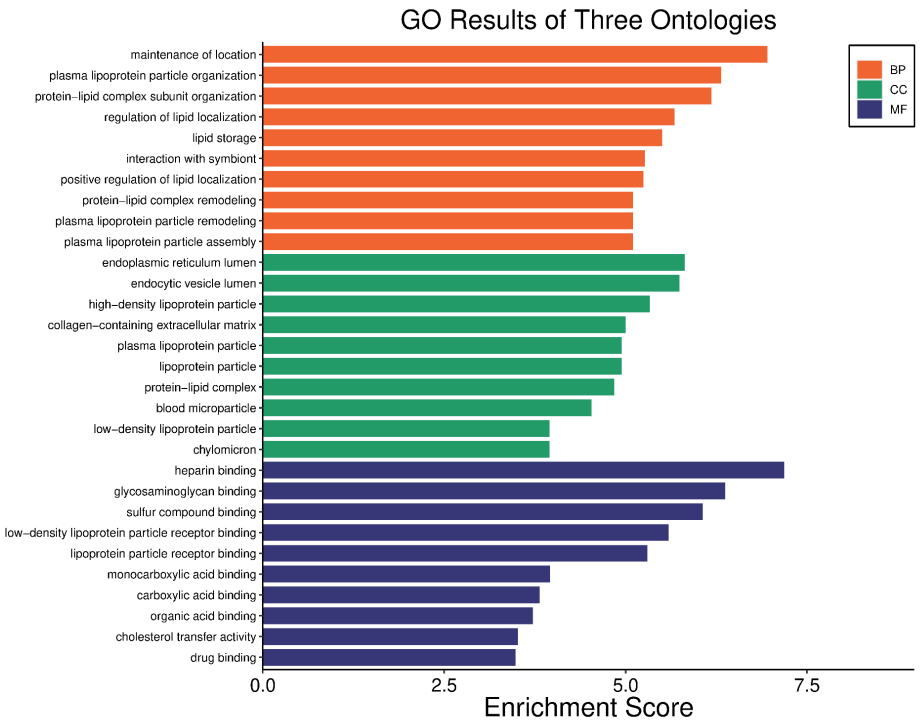

Figure 4a. This image represents the enrichment in all the ontologies from the GO results. BP stands for biological processes, CC stands for cellular components, and MF stands for molecular function. Four pathways were selected based on their high enrichment score. In BP, the

maintenance of location and plasma lipoprotein particle organization had high enrichment scores. In MF, the heparin and glycosaminoglycan bindings both had high enrichment scores as well. These processes were further analyzed with the genes due to the fact they had the highest enrichment out of all the results.

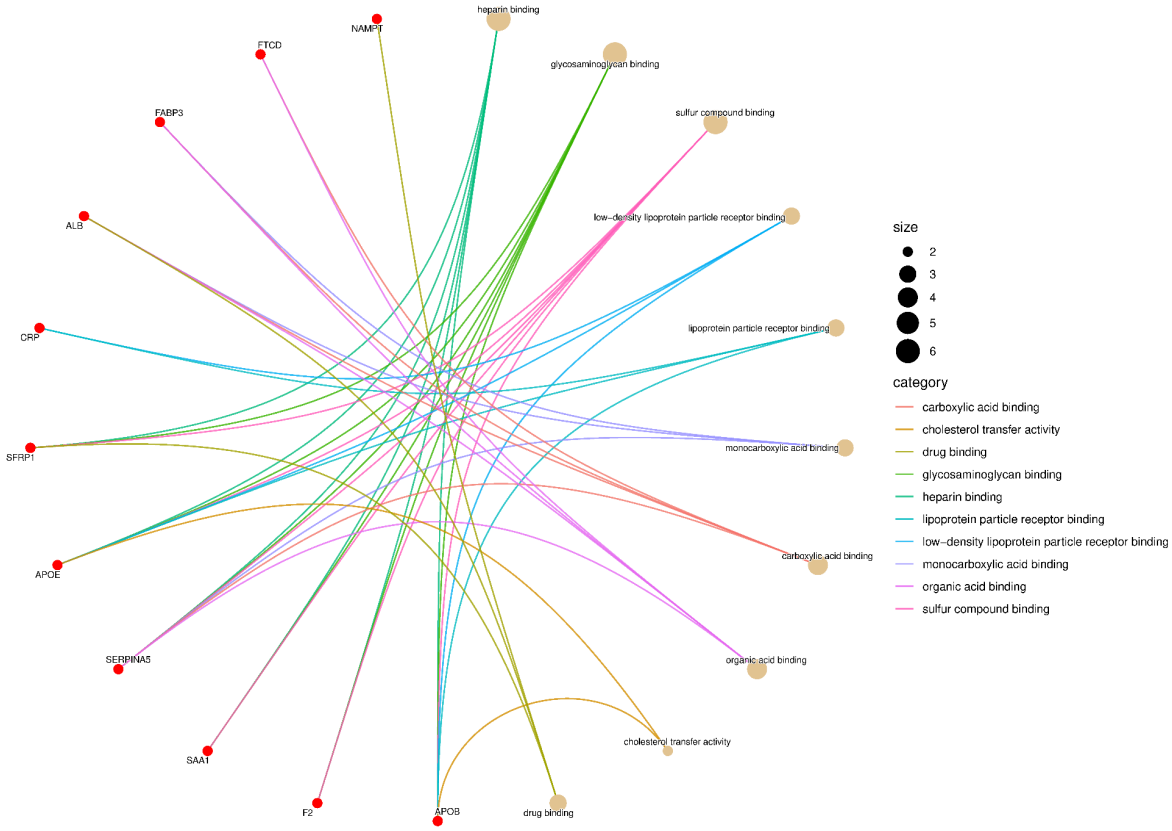

Figure 4b. This is a Cnet plot that shows which genes are related to certain pathways. We are looking for genes that are related to pathways that have a high enrichment score. APOE, F2, and APOB were genes that were related to glycosaminoglycan binding and heparin binding. These genes were specifically selected from other genes also related to glycosaminoglycan binding and heparin binding because they had similarities with the KEGG results.

Table 1. Summary of Key Genes Identified and their Connection to Alzheimer's Disease

Gene	**Pathway**	**Function and Connection with Broken Heart Syndrome**
APOE	Cholesterol Metabolism	The APOE gene is important in regulating cholesterol
F2	Complement and Coagulation	Produces prothrombin, a protein

	Cascade	that helps form blood clots
APOB	Cholesterol Metabolism	The APOB gene plays an important role in cholesterol metabolism

DISCUSSION

Summary of Findings

The goal of this research study was to identify genes that were significant in Stress-Induced Cardiomyopathy to understand the causes of the disease better. In the study, we found that complement and coagulation cascade pathways and cholesterol metabolism pathways were found to be significant. Also, the genes were all equally enriched in all three ontologies, but we found that MF and BP had the highest enrichments (Figure 4a). The top three significant genes that were prominent in both the KEGG and GO results were the APOE, F2, and APOB genes.

Interpretation of Results

From the KEGG results, we found that coagulation complement cascades and cholesterol metabolism are very important factors. This can help show that in addition to stress, high cholesterol, and blood clotting can also be important factors in causing Stress-Induced Cardiomyopathy. The analysis also found that the highest enriched ontologies were in MF and BP, which helps us know that causes can most likely be related to biological processes and molecular functions. Out of the genes that were present in the KEGG and GO results, the APOE, F2, AND APOB genes were significant in both results. This could mean that these genes are the most significant in Stress-Induced Cardiomyopathy. We can use these genes to understand their activation, and how it can prevent Stress-Induced Cardiomyopathy.

Comparison with Previous Studies

In other studies, patients with Stress-Induced Cardiomyopathy have been found to have higher cholesterol levels (10). They believe that cholesterol could also be an influence in the disease. Genes that are related in regulating cholesterol metabolism are the APOE and APOB gene (11-12). Coagulation and complement cascades are often involved with blood clotting. In another study patients with Stress-Induced Cardiomyopathy were found to have significantly higher levels of coagulation and complements than a control group of healthy individuals (2). A gene that is associated with blood clotting is the F2 gene because it creates prothrombin, which results in blood clots (13).

Implications

Given the information on the genes involved in the cholesterol metabolism pathway and the complement and coagulation cascade pathway, we can use this information to determine if these genes are directly related to Stress-Induced Cardiomyopathy. If they are then we can observe the gene expression to see who may be at risk.

Limitations

One limitation of this study is that it may be harder to observe these identified genes in someone who has Stress-Induced Cardiomyopathy because it is such a rare disease. Given its rarity, it may be hard to find patients with the disease to experiment with. Another limitation is that the identified genes will need to be further studied in the laboratory or clinical environment before developing a cure.

Future Direction

The identified genes can be tested in the laboratory by scientists in the laboratory or clinical trials to determine if the genes are the most significant in causing Stress-Induced Cardiomyopathy and how they can be used for early diagnosis.

REFERENCES

1) Takotsubo Syndrome in Intensive Cardiac Care Unit: Challenges in Diagnosis and Management. Secondary Research. Rocco Antonio Montone, Giulia La Vecchia, Marco Giuseppe Del Buono, Antonio Abbate, Tommaso Sanna, Daniela Pedicino, Giampaolo Niccoli, Massimo Antonelli, Filippo Crea. 2022. ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0146280621002929.
2) Elias Rawish, Thomas Stiermaier, Francesco Santoro, Natale D. Brunetti, and Ingo Eitel. Secondary Research. 2021. Journal of Cardiovascular Development and Disease. https://www.mdpi.com/2077-0383/10/3/479.
3) Stress Cardiomyopathy: Is it limited to Takotsubo syndrome? Problems of definition. Secondary Research. Petr A. Sarapultsev, Alexey P. Sarapultsev. 2016. ScieneDirect. https://www.sciencedirect.com/science/article/abs/pii/S0167527316313882.
4) Update of Takotsubo Cardiomyopathy: Present Experience and Outlook for the Future. Secondary Research. Anastasiia V. Bairashevskaia, Sofiya Y. Belogurova, Mikhail R. Kondratiuk, Daria S. Rudnova, Susanna S. Sologova, Olga I. Tereshkina, Esma I. Avakyan. 2022. ScienceDirect. https://www.sciencedirect.com/science/article/pii/S2352906722000392
5) The Challenge of Takotsubo Syndrome: Heterogeneity of Clinical Features. Secondary Research. SchlossbauerSusanne A.,GhadriJelena-Rima,ScherffFrank,TemplinChristian. 2017. Swiss Medical Weekly. https://smw.ch/index.php/smw/article/view/2372/3627.
6) NCBI GEO: archive for functional genomics data sets, Secondary Research, Tanya Barrett, Stephen E. Wilhite, Pierre Ledoux, Carlos Evangelista, Irene F. Kim, Maxim Tomashevsky, Kimberly A. Marshall, Katherine H. Phillippy, Patti M. Sherman, Michelle Holko, Andrey Yefanov, Hyeseung Lee, Naigong Zhang, Cynthia L. Robertson, Nadezhda Serova, Sean Davis, Alexandra Soboleva, 2012, Nucleic Acids Research https://academic.oup.com/nar/article/41/D1/D991/1067995
7) The KEGG Database, Secondary Research, Minoru Kanehisa, Gregory bock, Jamie A. Goode, 2002, Wiley Online Library https://onlinelibrary.wiley.com/doi/abs/10.1002/0470857897.ch8
8) The what, where, how and why of gene ontology - a primer for bioinformatics, Secondary Research, Louis du Plessis, Nives Škunca, Christophe Dessimoz 2011 Briefings in Bioinformatics https://academic.oup.com/bib/article/12/6/723/221815
9) SRplot: A free online platform for data visualization and graphing, Secondary Research Doudou Tang,Mingjie Chen,Xinhua Huang,Guicheng Zhang,Lin Zeng,Guangsen Zhang,Shangjie Wu,Yewei Wang, 2023, PLOS One https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0294236

10) Liu, W., Ruiz-Velasco, A., Wang, S. et al. Metabolic stress-induced cardiomyopathy is caused by mitochondrial dysfunction due to attenuated Erk5 signaling. Nat Commun 8, 494 (2017). https://doi.org/10.1038/s41467-017-00664-8
11) HAGBERG, J. M., WILUND, K. R., & FERRELL, R. E. (2000). APO E gene and gene-environment effects on plasma lipoprotein-lipid levels. Physiological Genomics. https://doi.org/10.1152/physiolgenomics.2000.4.2.101
12) National Library of Medicine. (n.d.). F2 gene: Coagulation factor II, thrombin. MedlinePlus. Retrieved March 2, 2025, from https://medlineplus.gov/genetics/gene/f2/
13) Amanda J Whitfield, P Hugh R Barrett, Frank M van Bockxmeer, John R Burnett, Lipid Disorders and Mutations in the APOB Gene, Clinical Chemistry, Volume 50, Issue 10, 1 October 2004, Pages 1725–1732, https://doi.org/10.1373/clinchem.2004.038026