Edanur ERSÖZ

MASTER'S DEGREE THESIS

2025-ANTALYA

T.C.

AKDENIZ UNIVERSITY

THE INSTITUTE OF HEALTH SCIENCES

GENE AND CELL THERAPY DEPARTMENT

CORRELATION BETWEEN SAFETY ISSUES AND SUCCESS IN CLINICAL TRIALS OF ADVANCED THERAPY MEDICINAL PRODUCTS

Edanur ERSÖZ

MASTER'S DEGREE THESIS

2025-ANTALYA

T.C.
AKDENIZ UNIVERSITY
THE INSTITUTE OF HEALTH SCIENCES
GENE AND CELL THERAPY DEPARTMENT

CORRELATION BETWEEN SAFETY ISSUES AND SUCCESS IN CLINICAL TRIALS OF ADVANCED THERAPY MEDICINAL PRODUCTS

Edanur ERSÖZ

MASTER'S DEGREE THESIS

Advisor
Assoc. Prof. Dr. Devrim DEMİR DORA

"My thesis can be utilized by citing references"

2025-ANTALYA

To the Institute of Health Sciences;

This study has been approved by our jury as a Master of Science Thesis in the Department of Gene and Cell Therapy, Gene and Cell Therapy Master of Science program.

Signatures

Thesis Advisor : Assoc. Prof. Dr. Devrim DEMİR DORA
Akdeniz University

Member : Prof. Dr. Salih ŞANLIOĞLU
Akdeniz University

Member : Asst. Prof. Dr. Ongun Mehmet SAKA
Ankara University

This thesis has been deemed by the jury appointed by the Administrative Board of the Institute and accepted by the Administrative Board of the Institute with decision dated ……/……./……. and numbered ……/……..

Institue Director

DECLARATION

I affirm that this thesis is the result of my individual effort, free from any unethical conduct throughout the entire process, from planning to completion. I have obtained all information included in this thesis in compliance with ethical principles, and I have duly credited all external information and opinions that were not generated within the scope of this thesis. Furthermore, I have appropriately documented and listed all such references in the bibliography.

Edanur ERSÖZ

Thesis Advisor
Assoc. Prof. Dr. Devrim DEMİR DORA

ACKNOWLEDGEMENTS

As a woman striving to make her mark in the world of science, I would like to begin by expressing my deepest gratitude to Mustafa Kemal Atatürk, whose vision and reforms paved the way for women like me to pursue education and academic careers freely and proudly.

I extend my sincere thanks to Assoc. Prof. Dr. Devrim Demir Dora, whose guidance, encouragement, and unwavering support have been invaluable throughout my graduate studies. Being her student has been both an honour and a privilege.

I am also grateful to Prof. Dr. Salih Şanlıoğlu, whose insightful comments and encouraging academic contributions have inspired and supported my development during this journey.

I express my sincere thanks to the esteemed members of my thesis defence committee, Prof. Dr. Salih Şanlıoğlu, Assoc. Prof. Dr. Devrim Demir Dora, and Asst. Prof. Dr. Ongun Mehmet Saka.

To my beloved family, first and foremost my dear father İskender Ersöz, my dear mother Seval Ersöz, who have stood beside me with unwavering strength and belief, and who have done everything in their power to help me move forward with confidence—I owe you everything.

My heartfelt thanks go to Ahmet Ozan Özgen for his continuous academic and emotional support, and for always believing in me.

Lastly, I would like to thank my dear friends, whose presence and encouragement have been a source of strength and comfort throughout this challenging process.

ÖZET

Amaç: Bu çalışma, sonlandırılmış klinik araştırmalarda ileri tedavi tıbbi ürünlerine ilişkin advers ilaç reaksiyonlarını (ADR) ve bu reaksiyonların terapötik endikasyonlarla ilişkilerini analiz ederek bu ürünlerin güvenlilik profillerini araştırmayı amaçlamaktadır.

Yöntem: ClinicalTrials.gov üzerinden ileri tedavi tıbbi ürünlerini içeren sonlandırılmış klinik araştırmalar uygun anahtar kelimeler ve filtreler kullanılarak taranmıştır. Araştırmalar incelenmiş ve çalışılan ürünlerin gerçekten bir ileri tedavi tıbbi ürünü olup olmadığı doğrulanmıştır. Uygun çalışmalardan ADR ve endikasyon verileri derlenmiş, advers etkiler ile endikasyonlar arasındaki ilişkiyi değerlendirmek için Random Forest analizi uygulanmıştır.

Bulgular: 01.01.2014-31.12.2024 tarih aralığı için yapılan aramada 12.668 klinik araştırma arasından, ilgili filtremeler sonucunda 1512 sonlandırılmış çalışma elde edilmiştir. Bu çalışmalara ilişkin herbir araştırma ürünü kontrol edilmiş, ve ileri tedavi tıbbi ürünü ilintili olmayan çalışmalar elenmiştir. 294'ünün ileri tedavi tıbbi ürünü içerdiği tespit edilmiştir. Dışlama kriterlerinin uygulanmasının ardından 12 çalışmaya ilişkin 1062 advers reaksiyonun analizi kapsamına alınmıştır. Yaygın olarak bildirilen advers reaksiyonlar arasında sitokin salınım sendromu (40/1062, %3.8) bulantı (33/1062, %3.1) ve ölüm (31/1063, %2.9) yer almıştır.

Sonuç: Elde edilen bulgular, ileri tedavi tıbbi ürünlerinin ciddi ve özgün advers reaksiyon profillerine sahip olduğunu ve bu nedenle özel farmakovijilans stratejilerine ihtiyaç duyulduğunu göstermektedir. Özellikle immün yanıtla ilişkili advers etkilerin yüksek sıklıkta görülmesi, bu ürünlerin güvenlilik izlemlerinin daha etkin bir şekilde planlanması gerektiğini ortaya koymaktadır. Random Forest gibi makine öğrenmesi yöntemleri, ADR–endikasyon ilişkilerinin anlaşılmasına önemli katkı sağlayarak bu yenilikçi tedavilere yönelik fayda-risk değerlendirmesinin optimize edilmesine yardımcı olmuştur.

Anahtar Kelimeler: Gen tedavisi, hücre tedavisi, güvenlilik, advers ilaç reaksiyonu, klinik çalışma

ABSTRACT

Objective: This study aims to investigate the safety profiles of gene and cell therapy products by analysing adverse drug reactions (ADRs) and their associations with therapeutic indications in terminated clinical trials.

Method: Terminated clinical trials involving advanced therapy medicinal products were retrieved from ClinicalTrials.gov using appropriate keywords and filters. The studies were manually reviewed to verify whether the investigational products were genuinely gene or cell therapies. Relevant ADR and indication data were extracted from eligible trials, and a random forest analysis was applied to evaluate the relationships between adverse events and therapeutic indications.

Results: In a search conducted for the period between January 1, 2014, and December 31, 2024, 1512 completed studies were obtained from 12668 clinical trials as a result of relevant filtering. Each research product related to these studies was checked, and studies unrelated to advanced therapy medicinal products were excluded. It was determined that 294 of these studies included advanced therapy medicinal products. After applying the exclusion criteria, 1062 adverse reactions related to 12 studies were included in the analysis. The most commonly reported adverse reactions included cytokine release syndrome (40/1062, 3.8%), nausea (33/1062, 3.1%), and death (31/1063, 2.9%).

Conclusion: The findings demonstrate that advanced therapy medicinal product possess distinct and serious adverse reaction profiles, necessitating tailored pharmacovigilance strategies. The high incidence of immune-related adverse events emphasizes the need for more robust safety monitoring for these products. Machine learning methods such as random forest provided valuable insights into ADR–indication relationships, contributing to the optimization of benefit–risk assessment for these innovative therapies.

Keywords: Gene therapy, cell therapy, safety, adverse drug reaction, clinical trial

TABLE OF CONTENTS

ÖZET i

ABSTRACT ii

TABLE OF CONTENTS iii

SYMBOLS AND ABBREVIATIONS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION 1

1.1. Background and Rationale 2

1.2. Objectives of the Study 2

2. GENERAL INFORMATION 3

2.1. Advanced Therapy Medicinal Products (ATMPs) 3

2.1.1. Definition and Classification of ATMPs 3

2.1.2. Current Regulatory Landscape for ATMPs 5

2.1.3. Gene Therapy Medicinal Products (GTMP) 6

2.1.4. Somatic-Cell Therapy Medicinal Products (SCTMP) 8

2.1.5. Tissue-Engineered Products (TEP) 12

2.1.6. General Structure of Clinical Trials 12

2.1.7. Requirements for Safety Management in Clinical Trials 13

2.1.8. Impact of Safety Issues on Clinical Outcomes 14

2.2. Bayesian Statistic 15

2.3. Machine Learning 16

2.3.1. Decision Trees 17

3. MATERIAL AND METHODS 20

3.1. System and Softwares 24

3.1.1. Working Environment 24

3.1.2. SPSS 24

3.1.3. Excel 24

3.1.4. R and R Studio 25

3.2. Programming Languages 26

3.2.1. R Packages and Open-Source Tools 26
3.3. Public Resources and Databases 27
3.3.1. Literature 27
3.3.2. Clinical Trials 27
3.4. Dataset 28
3.5. Bayesian Analysis 30
3.6. Evaluation Methodology 32
3.7. Machine Learning 33
4. RESULTS AND DISCUSSION 38
4.1. Bayesian Analysis 38
4.2. Random Forest 38
4.3. Analysis Results and Relationships Between Codes and Model 42
5. CONCLUSION 55
6. REFERENCES 60
APPENDIXES
APPENDIX-1.
RESUME

SYMBOLS AND ABBREVIATIONS

ADR : Adverse Drug Reaction

AIDS : Acquired Immune Deficiency Syndrome

AI : Artificial Intelligence

ALB: : Albumin

ALP : Alkaline Phosphatase

ALT : Alanine Aminotransferase

AML : Acute Myeloid Leukemia

APCs : Antigen-Presenting Cells

AST : Aspartate Aminotransferase

ATMP : Advanced Therapy Medicinal Product

BLA : Biologics License Application

CAR-T : Chimeric Antigen Receptor T-cell

CBER : Center for Biologics Evaluation and Research

CEA : Carcinoembryonic Antigen

CFR : Code of Federal Regulations

CKD : Chronic Kidney Disease

CRAN : Comprehensive R Archive Network

CRF **:** Chronic Renal Failure

CRS : Cytokine Release Syndrome

CTCs : Circulating Tumor Cells

DCT : Dendritic Cell Treatment

DBIL : Direct Bilirubin

DFS : Disease-free Survival

EMA : European Medicines Agency

EMT : Epithelial-Mesenchymal Transition

EU : European Union

FDA : Food and Drug Administration

FDCA : Federal Food, Drug, and Cosmetic Act

FIH : First-in-Human

GGT : Gamma-Glutamyl Transferase

GTMP : Gene Therapy Medicinal Product

GVHD : Graft-versus-Host Disease

Hb **:** Hemoglobin

HSCT : Hematopoietic Stem Cell Transplantation

ICANS : Immune effector Cell-Associated Neurotoxicity Syndrome

IND : Investigational New Drug

LDH : Lactate Dehydrogenase

MedDRA : Medical Dictionary for Regulatory Activities

M1 : Apple M1 Processor

MCMC : Markov Chain Monte Carlo

MSE: : Mean Squared Error

NADCs : Non-AIDS-Defining Cancers

NCT : National Clinical Trial

NIR : No Information Rate

NPV : Negative Predictive Value

OS : Overall survival

PHSA : Public Health Services Act

PPV : Positive Predictive Value

QoL : Quality of life

R : R Programming Language

RNA : Ribonucleic Acid

SCTMP : Somatic Cell Therapy Medicinal Product

SPSS : Statistical Package for the Social Sciences

TBIL : Total Bilirubin

TEP : Tissue-Engineered Product

TMMDA : Turkish Medicines and Medicinal Devices Agency

TN : True Negative

TP : True Positive

US : United States

VLS : Vascular Leak Syndrome

LIST OF FIGURES

Figure 2.1. EMA classification for ATMPs. 5

Figure 2.2. *In vivo* and *ex vivo* gene therapy steps. 7

Figure 2.3. Gene therapy vectors. 8

Figure 2.4. CAR-T cell therapy process. 10

Figure 2.5. Schematic representation of CAR-Natural Killer (NK) therapeutic approach. 11

Figure 2.6. Phase of clinical studies. 13

Figure 2.7. Discontinuation rates in clinical trials. 15

Figure 3.1. The scope of the study. 23

Figure 4.1. A possible example of one of the created decision trees. 38

Figure 4.2. Distribution of adverse events frequency for terminated clinical trials. 42

Figure 4.3. The evaluation of disease-related adverse events in terms of seriousness. 44

Figure 4.4. Termination reasons of ATMP related clinical trials. 48

Figure 4.5. Distribution of all ATMP- terminated clinical trials by phases. 49

Figure 4.6. Distribution of terminated clinical trials by phases-Safety related terminated studies. 49

Figure 4.7. Classification of the studies by methodology. 50

Figure 4.8. Classification of the viral based gene manipulation 50

LIST OF TABLES

Table 3.1. Terminated studies because of safety related causes. 21

Table 3.2. Sample of data set. 29

Table 4.1. Confusion matrix. 39

1. INTRODUCTION

Gene and cell therapies have increasingly gained attention as promising approaches for treating a wide variety of diseases, particularly in oncology, genetic disorders, and immune-related conditions. While their therapeutic potential is substantial, these treatments also bring forth notable challenges related to safety and efficacy especially concerning the profile of adverse drug reactions (ADRs). Gaining a clearer understanding of the link between specific ADRs and the disease areas in which these therapies are applied could be essential for enhancing both patient safety and clinical success (Demir Dora, 2021).

This thesis explores the adverse events associated with gene and cell therapies in clinical trials, with a particular focus on those studies that were terminated. The research employs advanced statistical tools, including random forest analysis, to uncover potential relationships between ADRs and treatment indications.

The analysis draws upon data from 1062 ADRs reported in clinical trials that were discontinued due to safety-related concerns. These trials encompass a wide spectrum of gene and cell therapy modalities, providing a comprehensive view of ADR trends across various therapeutic approaches and patient demographics.

The primary objectives of this thesis are to:

a. Determination and examination of the ADRs documented in gene and cell therapy techniques' concluded clinical trials.
b. Examination of how certain indications in these studies relate to ADRs
c. To assess the significance and strength of these relationships, apply random forest analysis.
d. The results of this study are meant to help guide clinical judgment, add to the expanding corpus of information on the safety profiles of gene and cell therapies, and offer suggestions for the design of next clinical trials. Furthermore, knowing these connections will help with risk management strategy development and clinical practice for patient monitoring.

1.1. Background and Rationale

In the scope of this study, clinical trials involving gene and cell therapies that have been discontinued or terminated, primally because of safety concerns, are included. A substantial and varied dataset was obtained by analysing 1062 ADRs from the safety-related terminated clinical trials. A thorough examination of ADR patterns across multiple indications is made possible by these trials, which reflect different kinds of gene and cell treatments used in numerous therapeutic fields. From early-phase trials to more advanced research, the trials chosen for this analysis cover a variety of clinical development stages and represent a broad spectrum of patient groups and treatment conditions.

Due to the parallelism between Türkiye and the EU, the scope of this thesis will mainly be based on the European Medicine Agency (EMA) and the Turkish guidelines adapted from the EMA. The processes of ATMPs in Türkiye have been designed in parallel with the EU. The classification criteria made by the EU are valid. The approval and control of the processes is the responsibility of the Turkish health authority "Turkish Medicines and Medical Devices Agency (TMMDA)".

1.2. Objectives of the Study

This study's main goals are to find and examine ADRs that have been reported in gene and cell therapy clinical studies that have terminated. By evaluating the incidence and severity of adverse events in relation to various advanced therapy types, it seeks to investigate the connections between certain ADRs and the indications for which these therapies have been administered. To assess and measure the importance of these correlations, the main option was using the Bayesian analysis. Unfortunately, because of the environment of the related dataset, it has been seen that other statistical approached should be preferred. At this point, the study has been transformed to the random forest analysis, which offers a strong statistical foundation for spotting trends and forecasting ADR hazards. As a result, the essential aim is helping to create better patient monitoring and risk management plans, which will improve the security of gene and cell therapies.

It is expected that this study would provide insightful information about the safety profiles of gene and cell therapies, along with suggestions for future trial designs and clinical practice. The study intends to enhance patient safety, direct clinical decision-making, and promote the development of gene and cell therapy applications in medicine by

comprehending the particular ADRs connected to these treatments and their corresponding indications.

2. GENERAL INFORMATION

2.1 Advanced Therapy Medicinal Products (ATMPs)

Because technology is developing so quickly, advanced therapy medicinal products (ATMPs) are a cutting-edge class of varied, research-focused biopharmaceuticals. Tissue-engineered products (TEPs), somatic cell therapy medicinal products (SCTMPs), gene therapy medicinal products (GTMPs), and combination products that combine tissues or cells with a device are all included in this category in line with the EMA guidelines (European Medicines Agency (EMA), 2023).

On the other hand, the subclassification in the US includes the two main product classes of cellular therapy and gene therapy (Iglesias-Lopez et al., 2019).

The certain that the processing of cell- and tissue-based products involves manipulation that changes their biological properties is necessary for them to be classified as advanced therapies in both the US and the European Union (EU). However, the US has a more broader classification approach, mainly considering whether they qualify as biologic medicines, whereas the EU emphasizes distinct regulatory frameworks and unambiguous differentiation between these items (Iglesias-Lopez et al., 2019).

2.1.1. Definition and Classification of ATMPs

GTMP, SCTMP, and TEP are the three primary categories into which the EMA classifies ATMPs. Additionally, under another name, known as combined ATMPs, products that contain a medical device with one or more of the other classes are listed (Figure 2.1.). This classification is broader under the U.S. Food and Drug Administration (FDA), where it is divided into two categories: cell therapy products and gene therapy products. The FDA prefers the phrase "cellular and gene therapy products (CGTs)" over the sometimes-used term "advanced therapy" found in EMA recommendations (FDA, U.S. Food& Drug Administration Cellular & Gene Therapy Products, 2024).

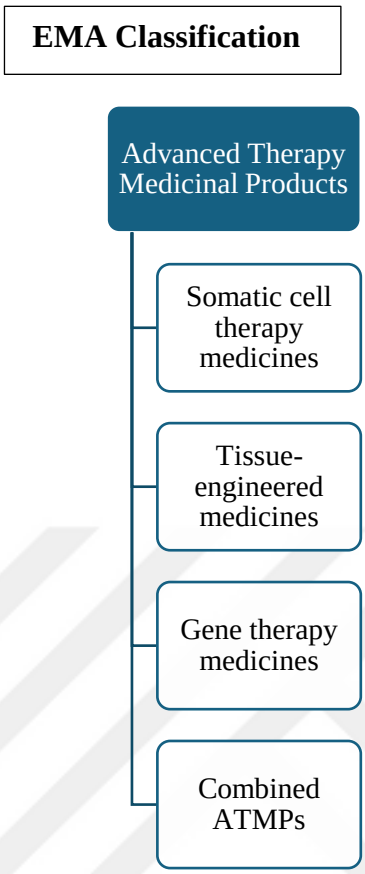


Figure 2.1. EMA classification for ATMPs.

2.1.2. Current Regulatory Landscape for ATMPs

In the EU, regulations for ATMPs were established through Regulation EC No. 1394/2007 by the EU Commission in 2007, with initial implementation starting in December 2008. The Committee for Advanced Therapies (CAT) of the EMA plays a significant role in the regulatory oversight of these products. CAT serves as the primary scientific committee responsible for evaluating marketing authorization (MA) applications for ATMPs and providing scientific recommendations for the classification of these products (Frederico et al., 2025).

The US regulatory framework, governed by the Federal Food, Drug, and Cosmetic Act (FDCA) and Public Health Services Act (PHSA), grants the FDA authority to regulate medicinal products. These acts control biological goods, including cutting-edge medicines. The Code of Federal Regulations (CFR) contains the FDA's regulations, with Title 21 offering special instructions for biological goods. For clinical studies, applicants must submit an Investigational New Drug (IND), and for marketing approval, they must

file a Biologics License Application (BLA). FDA centers are responsible for combination goods, whereas Center for Biologics Evaluation and Research (CBER) oversees biological products, including advanced therapies. In general, the framework guarantees the effectiveness and safety of pharmaceuticals for the general public's health. These differences highlight the diverse regulatory landscapes governing advanced therapies and emphasize the need for comprehensive evaluation criteria to ensure the safety and efficacy of these innovative treatments across borders (Iglesias-Lopez et al. 2019; Frederico et al. 2025).

2.1.3. Gene Therapy Medicinal Products (GTMP)

The EMA defines gene therapy as a method that works by introducing 'recombinant' genes into the body, typically to treat various conditions such as genetic disorders, cancer, or chronic diseases. These therapies have therapeutic, prophylactic, or diagnostic effects. A recombinant gene is a segment of DNA created in the laboratory by combining DNA from different sources (EMA, Advanced therapy medicinal products: Overview 2024).

In vivo and *ex vivo* approaches (Figure 2.2.) are the two main application categories for gene therapy techniques. The e*x vivo* method involves removing the patient's cells, genetically altering them, and then reintroducing them into the body. This technique exemplifies the somatic gene therapy strategy and has the benefit of using a variety of physical delivery mechanisms in addition to viral vectors, liposomes, and polymers. The *ex vivo* method works very well for improving cancer immunotherapy, treating infectious diseases like HIV, and treating blood disorders (Demir-Dora, 2024; Volodina and Smirnikhina, 2024)**.**

On the other hand, in *in vivo* gene therapy, genetic material is introduced directly into the body. Various delivery methods can be employed in this approach, including viral or non-viral methods. Each method is selected based on its suitability for the specific therapeutic target and its ability to efficiently deliver genetic material to the desired cells or tissues (Volodina and Smirnikhina, 2024).

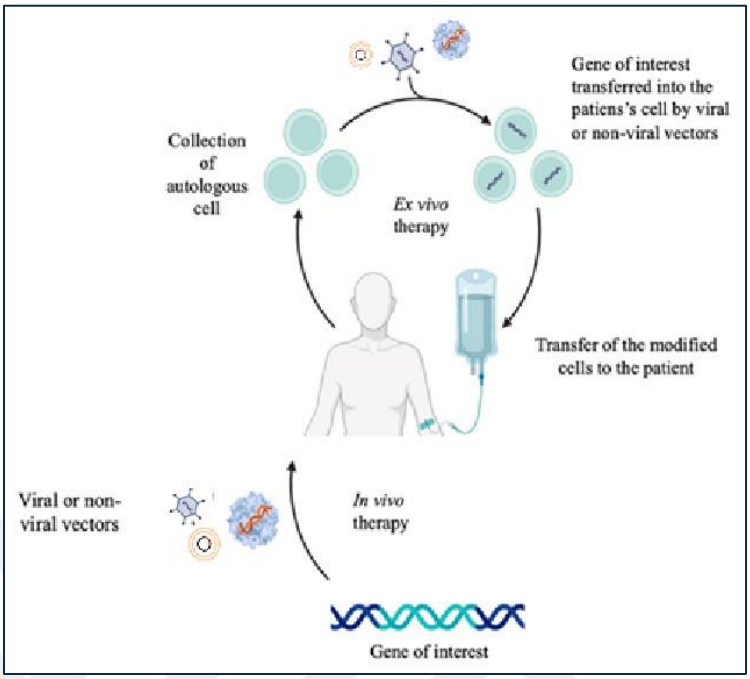


Figure 2.2. *In vivo* and *ex vivo* gene therapy steps.

Non-viral vectors are gaining attention as promising substitutes for viral vectors in gene delivery due to their advantageous safety features, such as reduced cytotoxicity, immunogenicity, and mutagenicity risk. A key benefit is their inability to trigger pre-existing immunity in the host, which substantially lowers the likelihood of immune-related adverse effects (Demir-Dora, 2021). However, these advantages are offset by notable limitations, including low gene transfer efficiency, inadequate targeting precision, short-lived gene expression, and generally limited therapeutic effectiveness (Pack et al., 2005). Various viral and non-viral vectors used in gene therapy are shown in Figure 2.3 (Butt et al., 2022).

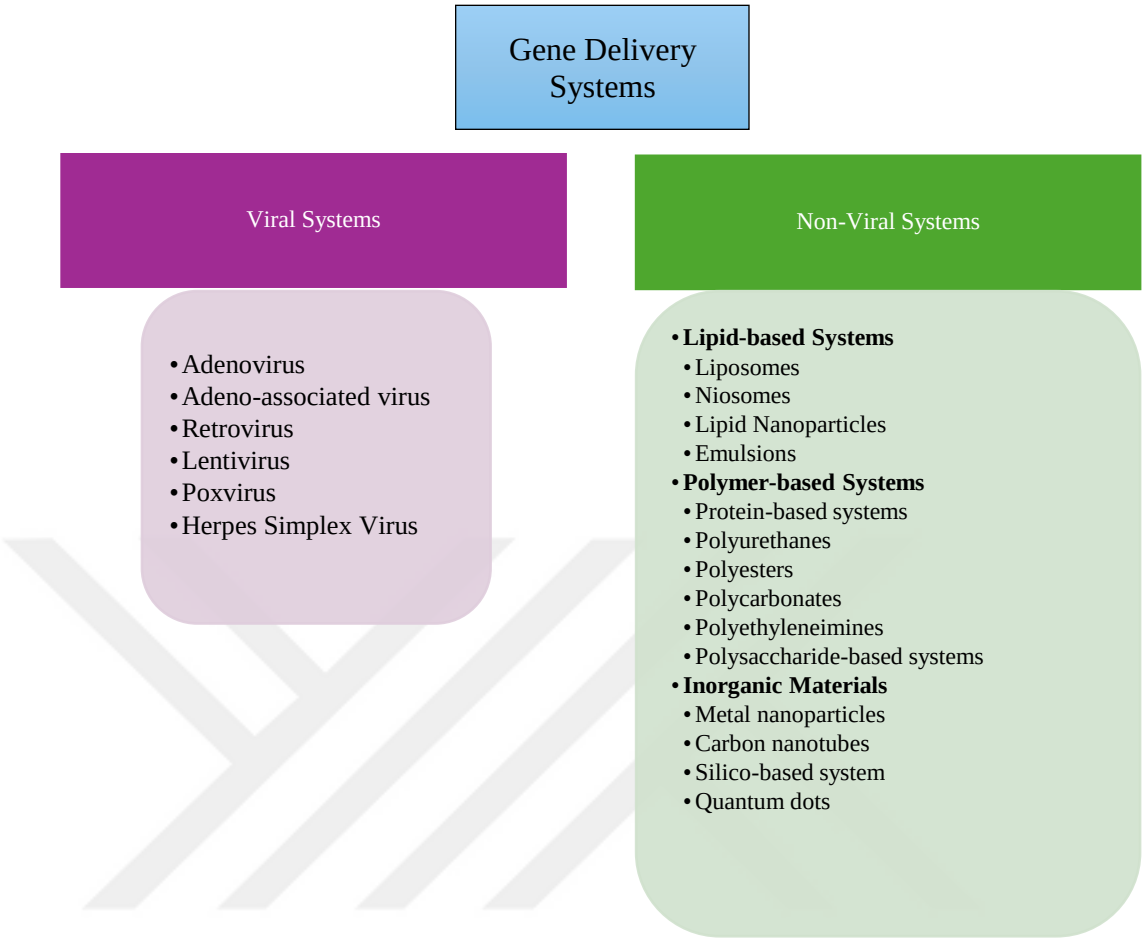


Figure 2.3. Gene delivery systems.

Gene therapy continues to advance with these diverse approaches, aiming to address a wide range of genetic and acquired diseases by targeting specific cells or tissues with therapeutic genes or gene-editing tools (Volodina and Smirnikhina, 2024).

2.1.4. Somatic-Cell Therapy Medicinal Products (SCTMP)

Somatic-cell therapy medicines consist of cells or tissues that have been modified to change their biological traits, or they may involve cells or tissues originally intended for different essential bodily functions. These therapies are utilized for treating, diagnosing, or preventing diseases (EMA, Advanced therapy medicinal products: Overview, 2024). Somatic cell therapy does not involve the modification of germ cells (sperm or egg cells), ensuring that genetic changes are confined to the treated individual and are not transmitted to subsequent generations. Various forms of somatic cell gene therapy have been explored for treating diseases not primarily driven by inherited genes, including AIDS and cancer (National Academies of Sciences et al., 2017).

Within the scope of somatic cell therapies, Chimeric Antigen Receptor (CAR) T-Cell Therapy holds a significant place in current clinical studies. As an approach to cancer immunotherapy, CAR T-cell therapy fundamentally involves the transplantation of a patient's own cells that have been modified to target cancer antigens (Sterner and Sterner, 2021). Each patient receives a unique treatment plan, which can be summed up as follows (Figure 2.4.): T-cells are extracted from the patient, reengineered in a lab to generate proteins known as chimeric antigen receptors (CARs) on their surface, and then the altered cells are returned to the patient. The treatment strategy works by means of the transformed cells' particular receptors attaching to particular proteins on the cancer cells, which kills the cells (Waldman et al., 2020). Although its novelty, CAR-T therapy has shown limitations such as a 60% patient relapse rate and serious side effects like cytokine release syndrome (CRS). These difficulties have led to a great deal of clinical research and efforts to improve the safety and effectiveness of CAR-T treatment. Chemotherapy integration and adjustments to the CAR-T administration schedule are two strategies being used to get over these challenges. Additionally, efforts are being made to produce next-generation CAR-T therapies, such as CAR-natural killer (CAR-NK) cells and CAR macrophages (CAR-Ms), to enhance therapeutic results and increase the range of applications for this cutting-edge therapy approach (Cesur-Ergün and Demir-Dora, 2023).

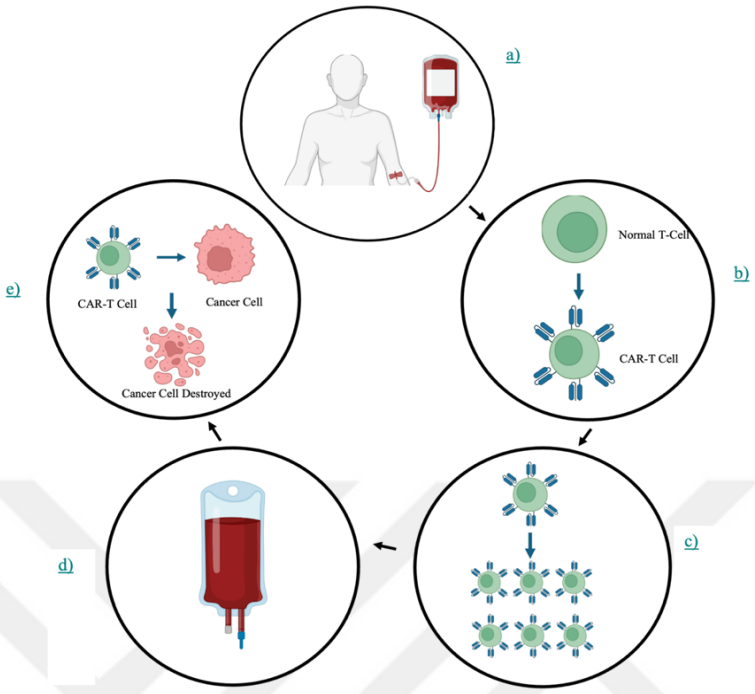


Figure 2.4. CAR-T cell therapy process.

a) Collection of T cells from the patient is performed.
b) T cells collected from the patient are genetically modified in the laboratory to express chimeric antigen receptors (CARs), which enable the engineered cells to recognize and eliminate cancer cells more effectively.
c) Expansion of the modified CAR-T cells in the laboratory environment is performed.
d) Following a short course of chemotherapy, CAR-T cells are infused back into the patient to target and destroy cancer cells.
e) CAR-T cells specifically find the cancer cells and destroy them.

Similarly, natural killer (NK) cell therapies involve analogous approaches. In the *ex vivo* approach, conditioned NK cells are expanded and activated in a laboratory setting from NK cells obtained from the patient or a healthy donor. Enhancing the anti-tumor activity of NK cells is achieved through the culturing of various growth factors and cytokines. These prepared NK cells are then injected into patients to more effectively recognize and kill tumor cells. The activity of NK cells can also be enhanced through genetic manipulation. CAR NK cells are produced by adding chimeric antigen receptors (CARs) to NK cells, enabling them to specifically bind to cancer cells (Figure 2.5.). CARs are designed to recognize specific antigens on the surface of cancer cells. This allows CAR NK cells to directly target and mount a more effective anti-tumor response. Currently,

there are no NK cell-based therapies approved by either the EMA or FDA (Vivier et al., 2024a).

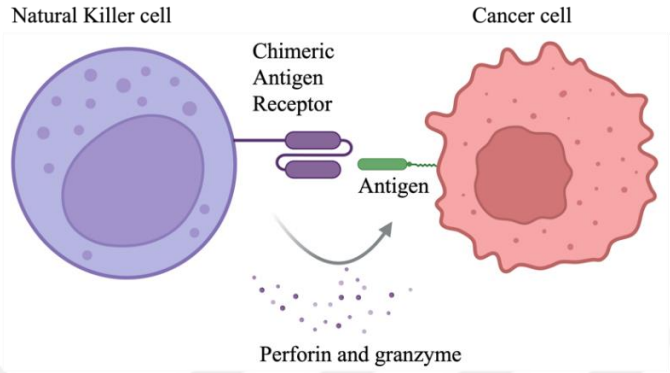


Figure 2.5. Schematic representation of CAR-Natural Killer (NK) therapeutic approach.

The alteration of dendritic cells, the immune system's most potent antigen-presenting cells (APCs), which can promote the innate immune response and trigger adaptive immunity, is the basis of dendritic cell treatment (DCT), another therapeutic strategy (Granato et al., 2024; Mantia-Smaldone and Chu, 2013). The procedure begins with the extraction of blood stem cells, which are subsequently cultured in a medium to develop into dendritic cells. The dendritic cells are exposed *in vitro* to tumor antigens, including tumor peptides, whole tumor cell lysates, nucleic acids encoding tumor antigens, apoptotic or necrotic tumor cells, tumor antigens and then injected into the patient (Benencia et al., 2012). Sensitized dendritic cells travel to the lymph nodes, where they induce the proliferation of specific lymphocytes. This allows the lymphocytes to attack cancer cells that contain the specific target antigen (Han and Wang, 2024).

Currently, there is one FDA-approved studies based on dendritic cell therapy which is Provenge (Sipuleucel-T) (Cheever and Higano, 2011).

2.1.5. Tissue-Engineered Products (TEP)

Tissue-engineered medicines include cells or tissues that have been altered to enable repair, regeneration, or replacement of human tissue. These advanced therapies are designed to facilitate the restoration of damaged or dysfunctional tissue in medical treatments (EMA, Advanced Therapy Medicinal Products: Overview, 2024).

2.1.6. General Structure of Clinical Trials

Phase 0 studies are preliminary investigations that involve minimal human exposure to the drug. These studies do not aim to achieve therapeutic or diagnostic outcomes, instead, they focus on exploratory objectives, such as screening potential candidates or conducting microdosing studies to gather early pharmacokinetic and pharmacodynamics data. Assessing safety and tolerability is typically the main goal of exploratory early-phase trials, especially for First-in-Human (FIH) investigations. Compared to other pharmaceuticals, investigational ATMPs frequently have special clinical safety concerns that must be considered while designing these trials. These worries could include the possibility of long-lasting or irreversible negative consequences, like delayed or long-term safety problems. Infections, immunogenicity or immunosuppression, ectopic tissue formation, genomic integration with gene therapy ATMPs, and the potential for malignant transformation are some examples of specific concerns (Kandi and Vadakedath, 2023).

Clinical trials undergo several stages to systematically assess a treatment's efficacy and safety. To evaluate the safety of the medication, Phase I trials usually include a small cohort of 10–40 patients. By testing the medication on 100–300 patients, Phase II trials broaden the focus and start assessing safety and efficacy. With 1000–3000 individuals, Phase III trials are larger and aim to monitor side effects, establish overall effectiveness, and validate the treatment's efficacy in comparison to conventional medicines. These trials can take several years to finish and frequently require cooperation from various institutions. Phase IV trials are carried out in real-world settings after a drug is approved by the authority to evaluate long-term side effects, confirm efficacy in clinical practice, and investigate further applications for various patient populations or illnesses. This staged strategy guarantees that the clinical efficacy of the intervention is determined in a way that is both safe and reproducible (Minneci and Deans, 2018). The phases of clinical trials are shown in Figure 2.6 (Kandi and Vadakedath, 2023).

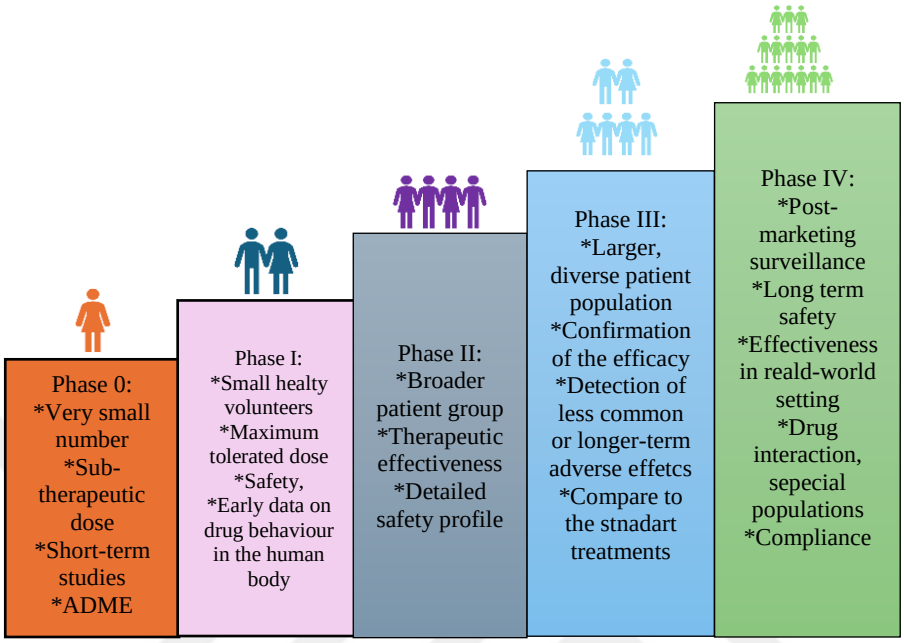


Figure 2.6. Phases of clinical trials.

2.1.7. Requirements for Safety Management in Clinical Trials

One of medicine's four core principles, "do no harm," serves as a cornerstone for the creation of novel therapeutic strategies. Clinical trials must adhere to safety management regulations starting in Phase 0, which must be monitored by the appropriate authorities and ensured by researchers (Kandi and Vadakedath, 2023).

Finding, evaluating, reducing, and effectively managing hazards is the aim of safety monitoring in clinical trials. Risk Management Plans, or RMPs, are required to be a part of the drug approval process in Europe by the EMA. A RMP offers a thorough synopsis of the major risks linked to the medication that have been discovered, as well as any potential hazards and information gaps. An action plan devoted to pharmacovigilance and risk reduction initiatives is based upon this overview. When preparing RMPs for ATMPs, defined risks, potential risks, and missing information specific to the technology of the ATMP should be addressed in detail (Yao et al., 2013).

The coordinator of the supporting research center is in charge of reporting any suspected unexpected serious adverse reactions (SUSARs) in accordance with Turkish legislation. They also have to make sure that the suspected product and the exact title of the study are

clearly defined. The investigational product's SUSARs, such as those from domestic or foreign studies carried out by the same sponsor, unplanned international reports, scientific literature, or regulatory bodies, must also be recorded. These must be compiled and reported at least every six months. Any safety concerns that could impact the benefit-risk assessment of the investigational product, require changes to its administration, or affect the study's design must also be reported. For example, inefficacy in a treatment for life-threatening diseases could pose significant risks to participants. Additionally, SUSARs involving the suspected product used in the study, especially if it is licensed in Türkiye, must be reported to the ethics committee and the Turkish Medicines and Medical Devices Agency (Klinik Araştırmalarda Güvenlilik Bildirimlerine İlişkin Kılavuz, 2024).

2.1.8. Impact of Safety Issues on Clinical Outcomes

Between 1999 and June 2015, a total of 143 clinical trials involving Advanced Therapy Medicinal Products (ATMPs) were either withdrawn, terminated, or ended according to the newest article. In contrast, 474 clinical trials, both ongoing and completed, were identified between 1999 and June 2013. This results in a discontinuation rate of 23.18% for ATMP trials, which is comparable to the rate for non-ATMP drug trials in development. The discontinuation rates for cell therapies, gene therapies, and TEP are 27.35%, 16.28%, and 16.34%, respectively. Oncology trials have the highest discontinuation rate at 43%, followed by cardiology at 19.2%. The rate of discontinuation is almost the same for both commercial and non-commercial sponsors, indicating that financial reasons may not be the primary cause for the discontinuation of these trials (Hanna et al., 2016a).

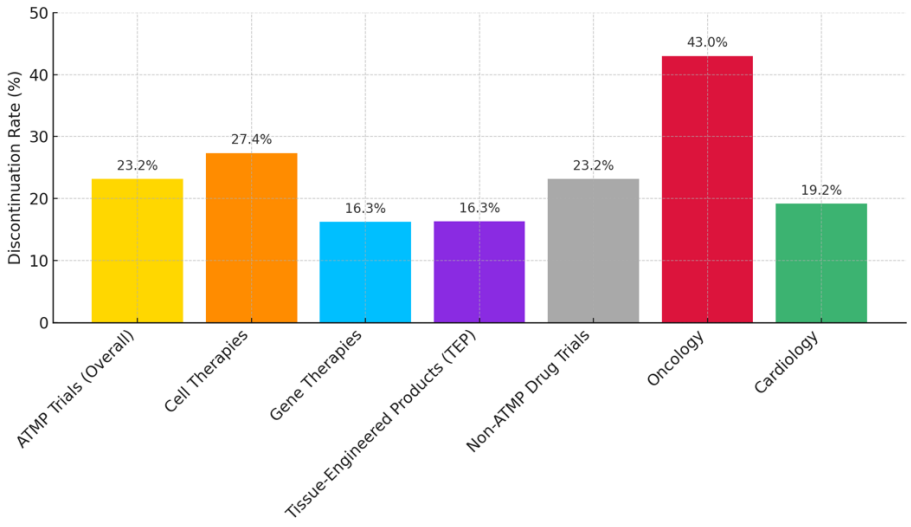


Figure 2.7. Discontinuation rates in clinical trials.

Hanna E. et al. analysed 71 systematic reviews, with 50 focusing primarily on effectiveness and safety. The findings suggest that while many ATMPs appear to have a favourable safety profile, significant uncertainties persist. The study emphasizes the necessity for collaborative research networks spanning *in-vitro*, preclinical, and clinical phases, and advocates for increased data sharing, particularly in preclinical stages and post-marketing real-world evidence collection. Additionally, it highlights the variability in how European Union (Aguilera-Cobos et al., 2022).

2.2. Bayesian Statistic

The Bayesian approach defines probability as the strength of a belief, unlike the classical (frequentist) approach to statistical inference, which treats probability as the long-term frequency of an event. This approach allows decision making by combining past information and new data. Bayes' theorem expresses the relationship between the previously known probability (prior) of a hypothesis, the observed data (likelihood) and the updated probability (posterior) in the light of new information. Clinical researches use Bayesian statistics to enhance decision-making and lower uncertainty (Lee, 2024).

The ability of the Bayesian technique to draw conclusions based on all available data and generate updated results is one of its most significant advantages. Because of this, it is possible to make significant predictions even when there are few observations.

Additionally, Bayesian model is adaptable and can be modified on a regular basis in response to new data. It is frequently employed in clinical research adaptable trial designs, enabling statistical judgments to be made in accordance with available data and therapeutic priorities (Muehlemann et al., 2023).

However, the Bayesian approach has also some limitations. One of the biggest challenge is that the choice of prior distribution may be subjective. There are different views on how the prior should be determined and this may lead different researchers to reach out different conclusions (Lee, 2024).

Compared to other conventional techniques, Bayesian statistics provides more adaptable and information-sensitive method. It is especially helpful in clinical studies if there are few observations or where there is a lot of prior knowledge. However, in certain situations, the method's application may be restricted by computing costs and uncertainty in earlier selection (Muehlemann et al., 2023).

2.3. Machine Learning

The automated process of finding patterns in data to carry out tasks like classification and prediction is known as machine learning (ML). ML is an area of artificial intelligence that includes a range of tools for data analysis and classification, or prediction based on patterns found. It functions under the more general heading of artificial intelligence, which aims to replicate intelligent computer processing and decision-making (Sarker, 2021)

Similar to conventional regression-based techniques, machine learning necessitates a number of crucial user-defined decisions. These include choosing the dataset to train the model, defining the variables that will be used for the classification or prediction process, and defining the outcome of interest for classification or prediction. Random forests and decision trees are two of the most popular supervised machine learning algorithms (Black et al., 2023).

2.3.1. Decision Trees

Decision trees are a dependable and efficient tool for making decisions because of their high classification accuracy and clear display of the gathered data. They are extensively utilized in a variety of domains, including marketing, finance, bioinformatics, and medical decision-making. Decision trees are helpful when examining intricate relationships between variables with accuracy and clarity. To construct progressively homogeneous (similar) groups with respect to the outcome, a decision tree divides the dataset into two or more subgroups in a sequential manner depending on predictive criteria. First, the most informative variable that maximizes homogeneity within the resultant groups is chosen. Until a predetermined stopping criterion—such as reaching a minimum number of observations per node, a maximum depth, or a particular level of purity—is satisfied, splitting keeps going recursively (Blockeel et al., 2023).

Various decision tree types exist depending on the target variable's characteristics:

a. Classification Trees: Used when the target variable is categorical. These trees classify data points into predefined categories, commonly assessed using metrics such as Gini impurity or entropy (information gain).
b. Regression Trees: Applied when there is a continuous target variable. By minimizing the variance within each split—often using mean squared error (MSE) as the criterion—these trees forecast numerical results.
c. CART (Classification and Regression Trees): An effective strategy that uses binary splits to produce an organized decision framework that supports both classification and regression models.
d. ID3 (Iterative Dichotomiser 3): An early algorithm that chooses the optimal characteristic at each split based on information gain.
e. C4.5: An extension of ID3 that uses a split ratio to normalize information gain and enhances handling of missing values and categorical data.
f. CHAID (Chi-squared Automatic Interaction Detection): This statistical method is helpful for categorical data analysis since it uses the chi-square test to generate multi-way splits rather than binary splits.

g. Random Forest: To improve accuracy and robustness, the Random Forest ensemble learning algorithm constructs several decision trees and aggregates their predictions. It was created by Leo Breiman in 2001 and introduces randomness into both feature selection and data sampling to overcome the main drawbacks of individual decision trees, including overfitting and noise sensitivity.

The Random Forest algorithm creates several decision trees that are each trained on a randomly selected sample of the training data using a technique known as bootstrap aggregating (bagging). Each tree is built using a unique set of attributes to guarantee model variety. The method creates predictions by combining the outputs of each tree once it has been constructed. For classification problems, the final prediction is decided by majority voting, whereas for regression tasks, the mean forecast of all trees is used.

The capacity of Random Forest to manage high-dimensional datasets with numerous variables is one of its main advantages. It lowers the model's variance by averaging several decision trees, which improves generalization on unknown data. The program also yields feature importance scores, which let researchers determine which factors in a dataset have the greatest influence. Because of this, Random Forest is especially useful in domains where knowing important predictors is essential, such bioinformatics, economics, and medical research.

Additionally, Random Forest has a significant degree of resistance to overfitting in comparison to single decision trees. Because each tree is trained on a different subset of data, the entire model is less vulnerable to specific data points that might act as outliers. Furthermore, by using only a random set of data for each split, Random Forest ensures that no single variable dominates the decision-making process, promoting a better-balanced model. On the other side, Random Forest has many drawbacks despite its great benefits. When working with very large datasets, the computational cost rises with the number of trees, making it slower than simpler models. Furthermore, Random Forest is less transparent than a single decision tree, where decision-making can be more readily traced, even though it improves interpretability over intricate deep learning models. The study offers a powerful and dependable method for categorizing adverse events and spotting

trends in gene and cell therapy clinical trials by utilizing Random Forest's advantages. The model is a great option for examining safety-related terminations since it can manage extensive feature connections while reducing overfitting (Cappelli et al., 2024).

3. MATERIAL AND METHODS

In this thesis study, the ClinicalTrials.gov website was searched for clinical trials that had been "terminated," "suspended," or "withdrawn" using particular keywords associated with advanced therapy medicinal products. The search was conducted using the following keywords: "gene therapy," "cell therapy," "CAR-T cell therapy," “tissue-engineered” and "somatic cell therapy." The exclusion of the duplication was carried out to remove duplicate entries and guarantee the dataset's accuracy after pertinent clinical trials were retrieved. The selection process was further refined by applying the necessary exclusions. Only the interventional” studies were selected.

In a search conducted for the period between January 1, 2014, and December 31, 2024, 1512 completed studies were obtained from 12668 clinical trials as a result of relevant filtering. At this point, every existed potential ATMP related studies were investigated upon their study product. To ensure that only studies with actual patient participation were considered, studies with enrolment status of at least one participant were selected.

Further examination of the grounds for discontinuance revealed that 17 clinical trials had been stopped because of claims of ineffectiveness, and 20 trials had been stopped because of safety concerns. In addition to safety and effectiveness concerns, sponsor decisions, difficulties enrolling patients, and budgetary limitations were often mentioned reasons for termination.

As only 12 studies of 20 studies provided open-sourced data related to safety concerns, so the analysis of adverse effects focused specifically on these 12 studies (Table 3.1.). Although the overall rate of termination due to safety concerns in gene and cell therapy trials was calculated to be approximately 6.2%, it is important to note that the total number of patients affected in these 12 safety-related terminations was recorded as 1062. This finding highlights the importance of monitoring and addressing safety challenges in the development and clinical application of gene and cell therapy products. The scope of the study was shown in Table 3.1.

Table 3.1. Terminated studies because of safety related causes.

Clinical Trial Number	Sponsor	Method	Vector	Indication	Route of Administration	Phase	Termination Reason
NCT00534703	Imperial College London	Gene therapy	AAV1.SERCA 2a	-Chronic Heart Failure	Percutaneous	Phase II	Early termination following Trial Steering Committee recommendation (Adverse events reported)
NCT00593645	Washington University School of Medicine	Cell therapy	Stem cell infusion	-Acute Myeloid Leukemia	Intravenous infusion	Phase II	Toxicities were worse than expected
NCT00858793	Universitäts klinikum Hamburg-Eppendorf	Gene and Cell Therapy	Gene (M87o)-modified (lentiviral vector based gene manipulation), CD34+ peripheral blood progenitor cells (PBSC)	-AIDS-related Lymphoma	Intravenous infusion	Phase I-II	A leukaemia case was reported in patient treated with a similar vector. For safety risk we stopped recruitment
NCT00924287	National Cancer Institute (NCI)	CAR-T Cell Therapy	CAR-T Cell Therapy-Retroviral vector based gene manipulation	-Metastatic Cancer	Intravenous infusion	Phase I-II	This study was terminated after the first patient treated on study died as a result of the treatment.
NCT01050764	Everett Meyer	Cell Therapy	Stem cell transplantation	-Leukemia, -Acute Chronic Myelogenous Leukemia - Myelodysplastic Syndrome (MDS) -Non-Hodgkin Lymphoma -Chronic Lymphocytic Leukemia -Acute - Myelogenous Leukemia -Acute Lymphoblastic Leukemia (ALL)	Intravenous infusion	Phase I-II	Safety
NCT01195480	University College, London	CAR-T Cell therapy	Donor-derived EBV-specific cytotoxic T-cells (EBV-CTL) transduced with the retroviral vector SFGalpha-CD19-CD3zeta	-Acute Lymphoblastic Leukemia	Intravenous infusion	Phase I-II	Due to lack of biological efficacy and CD19 CAR CTL persistence

NCT01236573	National Institutes of Health Clinical Center (CC)	Gene and Cell Therapy	Modified IL-12-Retroviral based gene manipulation	-Skin Cancer -Metastatic Melanoma	Intravenous infusion	Phase I-II	Unexpected toxicities, likely due to TIL/IL-12 & low % of durable responses.
NCT01273181	National Cancer Institute (NCI)	Gene and Cell Therapy	TCR-Gene Engineered Lymphocytes-Retroviral vector based manipulation	-Metastatic Cancer -Metastatic Renal Cancer -Metastatic Melanoma	Intravenous infusion	Phase I-II	Adverse events reported
NCT02340156	SynerGene Therapeutics, Inc.	Gene Therapy	Non-viral(nanoli_ poplex) vector based manipulation	-Recurrent Glioblastoma	Intravenous infusion	Phase II	Adverse events reported
NCT02525263	Prokidney	Cell Therapy	Autologus kidney cells	-Chronic Kidney Disease -Type 2 Diabetes	Percutaneous	Phase II	The single patient enrolled began haemodialysis treatments.
NCT02695160	Sangamo Therapeutics	Gene Therapy	Adeno-associated virus (AAV)-derived vectors.	-Hemophilia B	Intravenous infusion	Phase I	Adverse events
NCT03330834	Sun Yatsen University	CAR-T Cell Therapy	PD-L1 CAR-Lentiviral vector based gene manipulation	-Advanced Lung Cancer	Intravenous infusion	Phase I	Serious adverse events

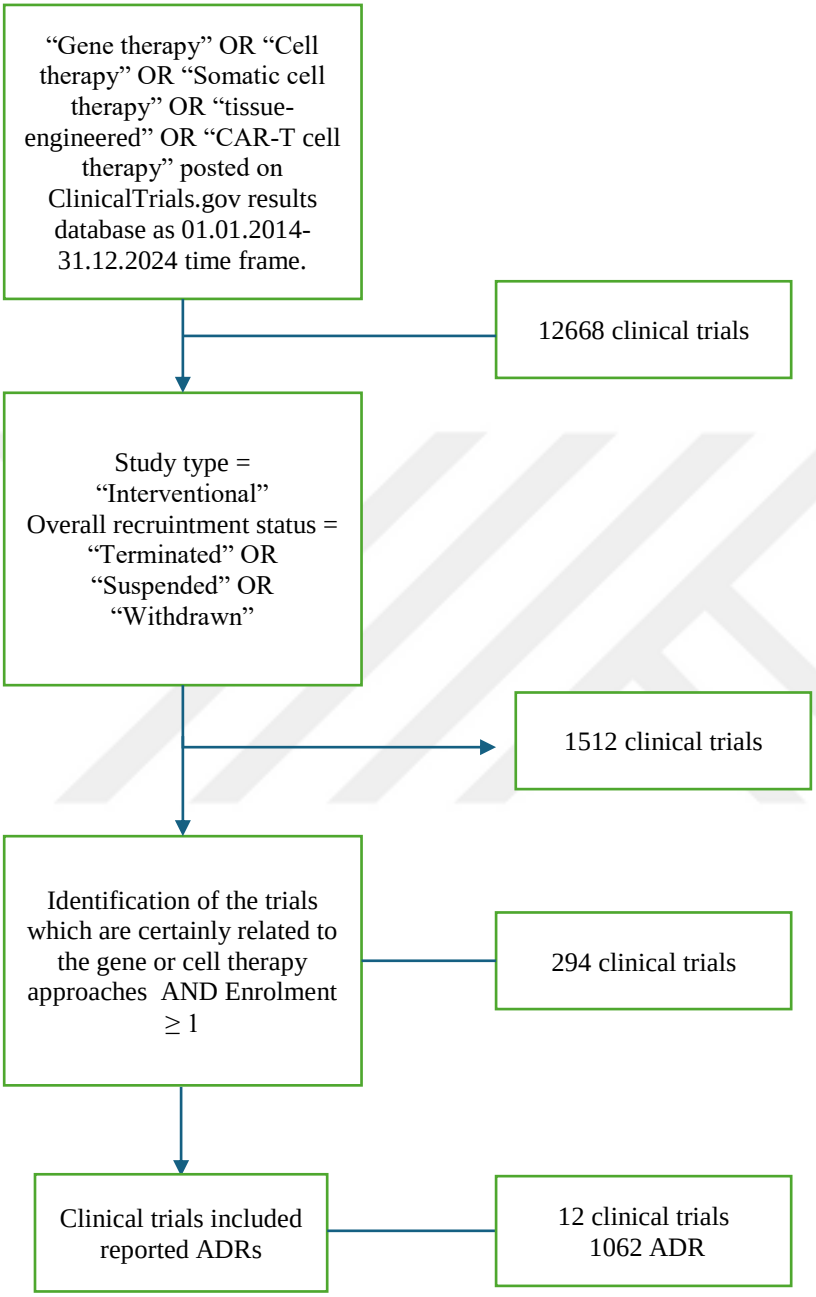


Figure 3.1. The scope of the study.

3.1. System and Softwares

3.1.1. Working Environment

This study was performed on a MacBook Air with M1 chip running macOS Sequoia Version 15.3.2. macOS is a Unix-based operating system developed by Apple Inc. and designed particularly for Mac computers.

3.1.2. SPSS

SPSS (Statistical Package for the Social Sciences) is a widely used software for statistical analysis. In this study, SPSS was used to perform basic statistical summaries on the data set, to perform data cleaning and to run various statistical tests. Moreover, its powerful graphical interface allowed for effective data visualization.

SPSS is particularly suitable for this study due to its user-friendly interface, comprehensive statistical tools and ability to efficiently process structured datasets. The software facilitated a systematic approach to data analysis, ensuring accuracy in descriptive statistics and hypothesis testing. Furthermore, its visualization features enabled clearer interpretation of the results, making it an important tool for discovering trends and patterns in the dataset.

The relationship between the “code” and “condition” parameters from the reviewed studies was examined using the McNemar test through SPSS. The McNemar test is formally defined as the classification of paired (or matched) responses in a binary (2x2) contingency table. In essence, it is a variant of the chi-square test specifically designed for dependent (i.e. related or paired) observations as opposed to independent (unrelated) samples. As a nonparametric statistical method, the McNemar test does not rely on normality assumptions and is therefore suitable for analyzing data sets that do not conform to a normal distribution (Adedokun and Burgess, 2012).

3.1.3. Excel

Excel is a spreadsheet program developed by Microsoft. It is used to organize data in tables, analyze it, create charts, and perform calculations. The most common uses include data entry and editing, mathematical and statistical calculations using formulas and functions, data analysis through filtering, sorting, and pivot tables, chart creation and visualization, and tasks related to reporting and budgeting. Each Excel file is called a

workbook, and it can contain multiple worksheets. Data is entered into cells, which are the intersections of rows and columns, such as cell A1.

In the examination of the dataset, Excel Version 16.92 (24120731) was used. The (range; criteria) formula was applied to determine the frequency of each adverse event reported in the dataset. The same approach was also utilized to identify the most frequently observed indication. Additionally, the charting features of Excel were used to visualize the findings derived from these analyses (Excel [Computer Software]. Microsoft Corporation., 2021).

3.1.4. R and R Studio

R version 4.4.0 and RStudio version "2024.12.1 Build 563" were used.

R Studio is an open source and free integrated development environment (IDE) for the R programming language. It provides an interface that facilitates the use of R's capabilities such as data analysis, statistical calculations and graphing. Thanks to its key features such as code editing, console, workspace management, graphing, debugging, package management, R Markdown, projects and version control, it is widely used in many fields such as data science, statistical analysis, econometrics, bioinformatics, social sciences and finance. With its user-friendly interface, comprehensive toolset and broad community support, R Studio is an indispensable tool for data analysis and statistical computing (Allaire, 2012).

R is an open-source programming language used for data visualization and statistical computation. It is a variation of the S programming language and was developed in the mid-1990s by Ross Ihaka and Robert Gentleman. It is preferred by a wide range of users because it is open source and free. It offers powerful functions for data analysis, hypothesis testing, regression models and time series analysis. It provides advanced data visualization capabilities with libraries such as ggplot2 and base graphics. Users can enhance R's functionality by downloading thousands of packages available on CRAN (Comprehensive R Archive Network). It runs on multiple operating systems, including Windows, macOS and Linux, making it widely accessible. Actively used by academics, data scientists and statisticians, R has a large online community. It is widely applied in

areas such as data science, big data analysis, machine learning, artificial intelligence, bioinformatics, genetic analysis, economics, finance and social sciences. Thanks to its powerful statistical analysis capabilities and data visualization tools, R has an important place in data-driven fields (Team, 2025).

3.2. Programming Languages

3.2.1. R Packages and Open-Source Tools

This study utilized several open-source tools and packages to facilitate data processing, statistical analysis, and machine learning modeling. The selection of these tools was based on their efficiency, flexibility, and compatibility with the methodologies employed in the research.

3.2.2. R Packages

The analysis was performed using several R packages, including dplyr, tidyr and brms, each serving a specific role in data processing and statistical modeling.

dplyr (1.1.4) is a powerful package designed for efficient data processing. It simplifies complex operations such as filtering, grouping, summarizing and transforming datasets, making it easier to handle large volumes of data. Its user-friendly syntax and optimized functions allow for fast and structured data editing, which is essential for preparing datasets before analysis (CRAN: Package dplyr).

tidyr (1.3.1.) plays a crucial role in reshaping and organizing data. It enables seamless conversion between long and wide data formats, which is particularly useful in structuring datasets for various statistical and machine learning models. In addition, tidyr includes functions to efficiently handle missing values and ensures that the dataset remains clean and properly formatted for subsequent analysis (CRAN: Package tidyr).

A package called brms (2.22.0) makes Bayesian statistical modeling easier in the R environment. brms, which is based on the Stan framework, makes it possible to create adaptable and incredibly adjustable Bayesian models. A large variety of probability distributions and model architectures are supported by this package, enabling researchers to draw reliable conclusions and integrate existing knowledge into their investigation.

Because of its sophisticated features, it's a useful tool for managing intricate statistical relationships and calculating data uncertainty (Bürkner, 2017).

3.3. Public Resources and Databases

ClinicalTrials.gov was used as the primary data source for the review of clinical trials related to gene and cell therapy. The dataset was filtered by specific keywords to identify trials classified as "Terminated", 'Suspended' or "Withdrawn". ADRs reported in these discontinued trials were analysed to assess their association with specific indications.

The articles reviewed in the literature search were accessed through PubMed and Google Scholar. Relevant scientific publications were reviewed to gather safety data, previous study findings and methodological information on ATMPs. A comprehensive literature review was conducted to assess the reasons behind early clinical trial terminations, evaluate adverse event profiles and improve the analytical approach of the study.

Guidelines and regulations issued by EMA, TMMDA and FDA were reviewed. Regulatory frameworks for ATMPs were reviewed to ensure that the study methodology and findings were in line with established safety and compliance standards. By integrating clinical data, scientific evidence and regulatory guidelines, the study provided a robust and well-validated analytical approach.

3.3.1. Literature

The literature published on ClinicalTrials.gov on terminated clinical trials was also reviewed to confirm the presence of any reported adverse events.

3.3.2. Clinical Trials

All completed gene and cell therapy trials were screened from ClinicalTrials.gov from 01.01.2014 to 31.12.2024. The grounds for ending the filtering trials were determined as described in Clinical Trials after a review of each one was conducted to confirm that they were, in fact, gene and cell therapy studies. The data set that was generated addressed clinical trials with safety-related concerns as the cause of the results individually, and the adverse effects found in the studies were assembled to the extent that they were evident in the data.

3.4. Dataset

Adverse event data from terminated safety-related clinical trials were obtained from ClinicalTrials.gov and included in the dataset with the relevant MedDRA-Preferred Term (PT) codes. PT codes are unique codes which are specific for each adverse event, universally. They were obtained from MedDRA (MedDRA-Help to Shape the MedDRA Terminology). Classification of adverse events as serious or non-serious was based on the investigator's judgement and recorded under the variable “severity criteria”. The therapeutic approach was divided into two main groups: “gene therapy” and “cell therapy”. Furthermore, the indication associated with each therapeutic approach was numerically coded from 1 to 12. A total of 1062 adverse events were evaluated in the dataset.

3.4.1. Characteristics and Statistics

The dataset consists of four columns and 1062 rows.

The first column, "method," represents a parameter that categorizes the investigated clinical study as either gene therapy or cell therapy. Gene therapy approaches are coded as "1," while cell therapy approaches are coded as "2".

The "condition" column represents the indication for which the therapeutic approach was used in the study. Each indication is assigned a number from 1 to 12, corresponding to the following conditions: 1: Acute Lymphoblastic Leukemia, 2: Advanced Lung Cancer, 3: AIDS-related Lymphoma, 4: Chronic Heart Failure, 5: Chronic Kidney Disease, 6: Hemophilia B, 7: Metastatic Cancer, 8: Metastatic Melanoma, 9: Multiple Myeloma, 10: Recurrent Glioblastoma, 11: Acute Myeloid Leukemia, and 12: Transplant Rejection.

Codes for adverse events recorded in clinical trials stopped due to safety concerns are listed in the third column. Each code is associated with a MedDRA PT code, which includes 138 different adverse events. PT codes are unique codes which are specific for each adverse event, universally. The severity criteria is shown in the fourth column. An adverse event is categorized as “1” when it is considered serious by the investigator and entered as “0” when it is considered non-serious. The related sample of columns were shown in Table 3.2.

Table 3.2. Sample of data set.

The main columns which were included in analysis was collored in gray.

Trial No.	Method	Method-Coded version	Condition	Condition-Coded version (1-12)	Adverse Event		MedDRA Code (PT)	Seriousness	Seriousness-Coded version
NCT 00534703	Gene Therapy	1	Chronic Heart Failure	4	Worsening heart failure		10066498	Serious	1
NCT 00534703	Gene Therapy	1	Chronic Heart Failure	4	Death		10011906	Serious	1
NCT 00534703	Gene Therapy	1	Chronic Heart Failure	4	LVAD Alarm		10052374	Serious	1
NCT 00534703	Gene Therapy	1	Chronic Heart Failure	4	Heart transplantation		10048769	Serious	1
NCT 00534703	Gene Therapy	1	Chronic Heart Failure	4	Iron deficiency-Anaemia		10002062	Non-serious	0
NCT 00534703	Gene Therapy	1	Chronic Heart Failure	4	Myocarditis		10028606	Non-serious	0

3.4.2. Feature and Target Variable Selection

The first analysis conducted selected "seriousness" as the target variable, with "method," "condition," and "code" as features. Additionally, manual analyses performed in Excel designated "condition" as the target variable, focusing primarily on the relationship between "code" and "condition," as well as examining associations with other parameters. Significant findings were highlighted properly.

3.5. Bayesian Analysis

Data analysis and Bayesian modelling were applied to a dataset to investigate the association between adverse event severity and other characteristics. Variables such as disease status, adverse event MedDRA code, severity level (serious or non-serious) and

treatment method (gene therapy or cell therapy) were all included in the dataset. Examining the distributions of the dataset, identifying missing data, cleaning the data and using the Bayesian model were all steps in the analysis process. Below is a comprehensive description of each of them:

a. Loading Libraries

The libraries used for analysis were dplyr, tidyr and brms:

dplyr: A library used for data manipulation, simplifying tasks such as grouping, filtering and summarizing.

tidyr: Used to convert data into long or wide formats, fill in missing values or prepare data for analysis.

brms: A package for Bayesian modelling in R, allowing the construction of flexible and robust models using the Stan framework.

b. Loading the Dataset

The data for analysis was loaded from a CSV file using the read.csv function in R. Users specified the file path and name to import the dataset into the R environment.

c. Examining Dataset Distributions

Frequencies of combinations for variables such as Method (therapy method), Condition (disease condition), code (MedDRA code), and Seriousness (seriousness level) were calculated:

The group by function was used to group data by the specified variables, ensuring unique combinations were grouped.

The summarise function was applied to calculate the total number of observations (Count) for each group, identifying unique combinations and their frequencies.

d. Analyzing Data Sparsity

To analyze sparsity, the data was reorganized into a “wide format.” This step aimed to represent each value of the Seriousness variable (e.g., serious or not serious) as separate columns. Using the pivot_wider function, missing observations were automatically filled with zeros (values_fill = 0). This format facilitated visualizing which seriousness levels were present for each combination.

e. Identifying Insufficient Data Groups

An evaluation threshold was applied to determine whether each combination had a sufficient number of observations for analysis. For instance, a minimum of five observations was required for a combination to be included in the analysis. The filter function identified groups falling below this threshold. This step ensured data adequacy for reliable results in Bayesian modeling.

f. Bayesian Modeling

A Bayesian model was built using Method, Condition, and Code to explain the Seriousness variable after the missing data was cleaned. The following was the modeling procedure:

Seriousness (seriousness level) is the dependent variable.

Method (cell therapy or gene therapy), Condition (disease state), and Code (MedDRA code) are independent variables.

Model structure: The brm function was used to construct a Bernoulli distribution model with a logit link.

Prior distributions: Each independent variable was given a normal (0, 1) prior distribution (prior = set_prior ("normal (0, 1)", class = "b").

Iterations and chains: Four chains and 1,000 iterations were used to run the model.

3.6. Evaluation Methodology

Data Quality and Preprocessing:

- Missing Data Handling: Evaluation of how missing values were addressed (imputation with mode for the Seriousness variable or removal) and how it impacted the model's ability to provide reliable predictions.
- Data Sparsity: Analysed the extent to which combinations of variables (therapy method, disease condition, and MedDRA code) had sufficient observations for valid statistical inference.
- Model Construction:
- Choice of Priors: The evaluation of prior distributions (normal (0, 1) priors for the independent variables) to ensure non-informative priors were used and that the model was not biased by prior assumptions.
- Model Structure: The decision to use a Bernoulli distribution with a logit link was assessed for its appropriateness in modelling the binary nature of the seriousness variable (serious vs. non-serious).
- Model Convergence and Diagnostics:
- Convergence: The analysis of MCMC convergence was crucial to verify that the chains had stabilized and the posterior estimates were reliable. This was evaluated through visual inspection of trace plots and the Gelman-Rubin statistic.
- Posterior Distribution: The evaluation of the posterior distributions and credible intervals of model parameters was conducted to assess the uncertainty and precision of the parameter estimates.
- Model Performance:
- Predictive Accuracy: Cross-validation was employed to evaluate the model's generalizability and predictive power. The performance of the model in replicating observed outcomes was checked to ensure its robustness.
- Sensitivity Analysis: The model's sensitivity to different prior distributions was evaluated to understand how robust the results were to changes in prior assumptions.

By evaluating these key parameters, the study assessed the validity of the Bayesian modeling process and determined whether the model could reliably explain the relationship between therapy methods and the seriousness of adverse events.

*The analysis, which was intended to be performed using Bayesian analysis, was insufficient in obtaining the desired results, the study was continued using a machine learning approach. Because of this, related general information about used approached defined in next sections.

3.7. Machine Learning

3.7.1. Random Forest

The correlations between independent variables (Method, Code, and Condition) and the severity of adverse outcomes are efficiently examined by the random forest model. These studies allow for the early detection of possible risks and can enhance risk management during clinical trials.

In the analysis, a binary classification problem was tackled using the Random Forest technique. Several criteria were used to assess the model's performance, accounting for the impact of class imbalance on predictive performance.

Code: Refers to the distinct codes that MedDRA (Medical Dictionary for Regulatory Activities) assigns to adverse effects. The systematic identification of adverse effects is made possible by these codes.

Condition: Represents for the disease linked to the adverse effect.

Data processing, model training, and performance assessment are among the phases that form up the code:

a. Loading Libraries

To provide tools for parallel processing and model training, the required R libraries are loaded, such as caret, randomForest, and doParallel. In order to manipulate data and use the random forest method, several libraries are necessary.

b. Data Loading and Cleaning

Missing values (NA) are eliminated from the dataset to leave only complete observations. Improved model accuracy is ensured by cleaning missing data. In order for the random forest method to properly comprehend the data, the code and seriousness variables are also transformed into categorical (factor) variables.

c. Splitting the Dataset into Training and Testing Sets

20% of the dataset is used for testing, and the remaining 80% is used for training. The testing set assesses the model's capacity for generalization, while the training set helps in the learning process. The Seriousness variable's class distribution is maintained during the splitting procedure, providing balanced subsets.

d. Enhancing Performance with Parallel Processing

Combining several decision trees is a computationally demanding step in the random forest technique. Model training is greatly accelerated by using parallel processing techniques to fully utilize the computer's CPU cores.

e. Training the Random Forest Model

During model training, Seriousness is the dependent variable, while all other features in the dataset serve as independent variables. The random forest algorithm constructs multiple decision trees and aggregates their results to perform classification. This approach is particularly effective for high-dimensional datasets and robust against overfitting.

f. Making Predictions

Following training, the model predicts the degree of severity of adverse effects on the test dataset. These predictions offer an objective evaluation of the model's capacity for generalization because the test dataset includes observations that have never been seen before.

g. Evaluating Model Performance

The model's performance is evaluated by comparing predictions with actual values in the test dataset. A confusion matrix is used to measure key metrics such as accuracy, sensitivity, and specificity. These metrics are crucial for assessing the model's ability to accurately predict serious adverse effects.

- Accuracy: The percentage of accurate classifications the model produced. Although a high accuracy suggests that the model functions well overall, it may not be accurate when there is a class imbalance.

- No Information Rate (NIR): The accuracy that the model might produce by merely predicting the most prevalent class, without offering any other helpful information. It is a sign that the model is meaningless if its accuracy is below this rate.

- P-Value [Acc > NIR]: This tests whether the model's accuracy is greater than the No Information Rate. A low p-value (usually < 0.05) indicates that the accuracy of the model is statistically significant and not due to chance.
- Kappa: Measures the agreement between the model's classifications and random classifications. A kappa value close to 1 indicates perfect agreement, while a value close to 0 indicates random predictions.

Sensitivity and Specificity:

- Sensitivity (Recall for class 0): The rate at which class 0 is correctly identified. A high sensitivity indicates that the model is good at recognizing class 0.
- Specificity (Recall for class 1): The rate at which class 1 is correctly identified. A high specificity indicates that the model is good at recognizing class 1.
- Positive Predictive Value (Precision for class 0): The proportion of positive predictions for class 0 that are actually correct. A high value indicates that the model's predictions for class 0 are accurate.

- Negative Predictive Value (Precision for class 1): The proportion of positive predictions for class 1 that are actually correct. A high value indicates that the model's predictions for class 1 are accurate.

Prevalence and Detection Rates:

- Prevalence: The proportion of class 0 in the dataset, representing the general distribution.
- Detection Rate: The proportion of actual class 0 instances that the model correctly identifies.
- Detection Prevalence: The proportion of predicted class 0 instances made by the model.

Balanced Accuracy: This evaluates the accuracy while considering the class imbalance. High balanced accuracy indicates that the model is correctly identifying both classes.

Interpretation of Possible Results:

- High Accuracy: Indicates that the model is making correct classifications, but it may be misleading due to class imbalance.
- Low Sensitivity: Suggests that the model is weak in identifying class 0.
- Low Specificity: Suggests that the model is weak in identifying class 1.
- Low Precision (Negative Predictive Value): Indicates that the model's predictions may be largely incorrect.
- High Balanced Accuracy: Reflects that the model is performing well in recognizing both classes.

4. RESULTS AND DISCUSSION

4.1. Bayesian Analysis

In this thesis, the dataset is assessed for the feasibility of conducting Bayesian analysis, which is commonly used for statistical inference in situations where assumptions and sufficient data are critical. The dataset includes 294 unique combinations of factors such as therapy method, condition, MedDRA code for side effects, and seriousness of side effects. The analysis highlights three primary issues that render Bayesian analysis infeasible:

Sparsity: The data is highly segmented across multiple variables, which leads to a significant number of groups with insufficient observations. With 210 out of 294 groups containing fewer than five observations, the data becomes too sparse to generate reliable posterior distributions using Bayesian methods. This sparsity reduces the statistical power required for accurate inferences.

Non-informative Priors: Given the limited data in several groups, the priors used in the Bayesian model would lack the necessary information to generate meaningful posterior distributions. As a result, the model would struggle to produce reliable estimates.

Covariance Structure: The relationships between variables in the dataset are too weak or non-existent. This is evident from the sparsity analysis, where many of the data combinations have very few observations, limiting the model's ability to capture meaningful relationships.

Several steps which are mentioned taken to analyze the sparsity, clean the dataset, and attempt a Bayesian model using the brms package. However, despite the data being cleaned and missing values imputed, the model fails to compile or converge. This further supports the argument that Bayesian analysis is not feasible with this dataset due to the structure and quality of the data.

4.2. Random Forest

The optimal number of decision trees enhances predictive capability. Figure 4.1 shows an possible example of one of the 500 decision trees created.

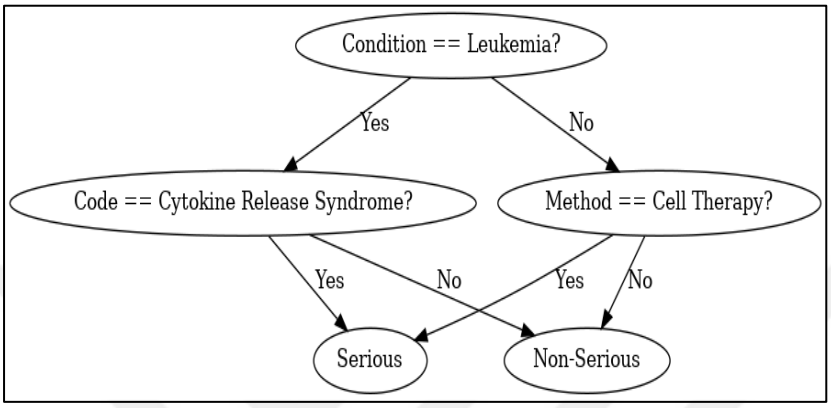


Figure 4.1. A possible example of one of the created decision trees.

The confusion matrix in Table 4.1. summarizes the model's predictions and their alignment with the actual class labels:

Table 4.1. Confusion matrix

Predicted/Actual	0	1
0	190	9
1	2	10

True Positives (TP): 190 cases were correctly classified as class 0.

False Positives (FP): 2 cases were incorrectly predicted as class 1, while they actually belonged to class 0.

False Negatives (FN): 9 cases were incorrectly predicted as class 0, while they actually belonged to class 1.

True Negatives (TN): 10 cases were correctly classified as class 1.

This confusion matrix serves as the foundation for calculating key performance metrics.

a. Accuracy and Statistical Significance

Accuracy:

The overall accuracy of the model was 94.79%, indicating that a high proportion of predictions were correct.

95% Confidence Interval: The accuracy is expected to fall between 90.86% and 97.37% under repeated sampling.

No Information Rate (NIR):

The NIR was 91%, which represents the accuracy that could be achieved by always predicting the majority class (class 0).

The model's accuracy exceeded the NIR, highlighting its capability to differentiate between the two classes.

P-Value [Accuracy > NIR]:

The p-value for the model's accuracy surpassing the NIR was 0.02887, which is statistically significant ($p < 0.05$). This confirms that the model performs better than random chance predictions.

Kappa Statistic:

The Kappa value was 0.6186, representing a moderate level of agreement between the model's predictions and actual labels, adjusted for chance agreement.

b. Sensitivity and Specificity

Sensitivity (Recall for Class 0):

The model correctly identified 98.96% of the instances belonging to class 0, demonstrating a high capacity for detecting the majority class.

Specificity (Recall for Class 1):

The model correctly identified 52.63% of the instances belonging to class 1, indicating a weaker ability to identify the minority class.

The disparity between sensitivity and specificity is indicative of class imbalance, which is common when one class dominates the dataset.

c. Predictive Values

Positive Predictive Value (PPV, Precision for Class 0):

When the model predicted class 0, it was correct 95.48% of the time.

Negative Predictive Value (NPV, Precision for Class 1):

When the model predicted class 1, it was correct 83.33% of the time.

These values highlight that the model is more reliable when predicting the majority class (0) than the minority class (1).

d. Prevalence and Detection Rates

Prevalence:

Class 0 comprised 91% of the dataset, while class 1 represented only 9%, illustrating a significant class imbalance.

Detection Rate:

The proportion of correctly predicted class 0 instances was 90.05%, close to the overall prevalence.

Detection Prevalence:

The proportion of instances predicted as class 0 was 94.31%, slightly higher than the actual prevalence, suggesting a bias towards the majority class.

e. Balanced Accuracy

The Balanced Accuracy, calculated as the average of sensitivity and specificity, was 75.79%.

This metric provides a more equitable evaluation of model performance by accounting for class imbalance.

f. Model Strengths and Weaknesses

Strengths:

The model demonstrates high overall accuracy (94.79%), making it effective at correctly classifying the majority of instances.

For class 0, the model shows excellent sensitivity (98.96%) and precision (95.48%), ensuring robust predictions for the dominant class.

Weaknesses:

The model struggles to accurately identify class 1, as evidenced by its lower specificity (52.63%) and moderate negative predictive value (83.33%).

The model's performance is noticeably skewed towards the majority class due to class imbalance.

For the majority class (0), the Random Forest model performs well, while for the minority class (1), it is less effective. The difficulties of dealing with uneven datasets are shown by this imbalance as well as the moderate agreement (Kappa = 0.6186).

Further refinement, such as addressing the class imbalance or exploring alternative algorithms, may enhance the model’s ability to generalize across both classes. Detailed attention to specificity and balanced accuracy will be key to improving its predictive capabilities.

4.3. Analysis Results and Relationships Between Codes and Model

The relationships between the conditions and MedDRA- PT codes identified within the scope of the thesis have been explained. The classification is based on McNemar test and

excel analysis result. In line with the Excel features, the most frequently occurring adverse events were identified, and the relationship between the detected adverse events and the indication and method was examined.

Within the scope of the study, the relationship and causality of the most frequently observed adverse reactions (Figure 4.2.) with the treatment indications were examined. At this point, nausea, vomiting and diarrhoea are among the side effects that are classified as non-serious but are frequently seen. Since they are among the side effects that are frequently seen even in conventional drugs, their relational evaluation was not prioritized. In addition to the adverse reactions listed the table, some adverse reactions that are considered clinically significant and require monitoring are also discussed. For some of these relationships, the adverse effects may naturally arise from the nature of the related disease, as compiled from literature studies and presented within the thesis.

The following part presents these relationships:

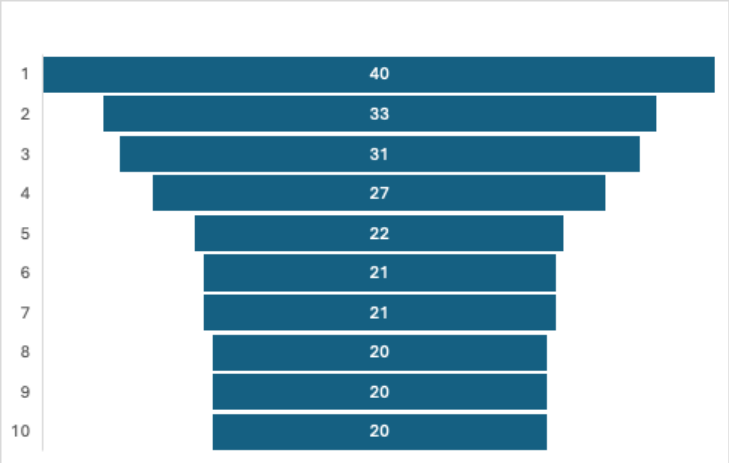


Figure 4.2. Distribution of adverse events frequency for terminated clinical trials.
1. 10052015- Cytokine release syndrome, 2. 10028813- Nausea, 3. 10011906- Death, 4. 10047700-Vomiting, 5. 10025327-Lymphopenia, 6. 10000167-Abnormal platelets, 7. 10016288- Febrile neutropenia, 8. 10010305-Confusional state, 9. 10012735- Diarrhoea, 10. 10021789-Infection

The adverse event coded as 10000167, abnormal platelets, has been associated with clinical studies involving gene therapy strategies targeting the metastatic cancer indication

and with clinical studies involving cell therapy strategies targeting the metastatic melanoma indication, based on the analysis conducted.

The adverse effect of alanine aminotransferase abnormal with code 10001547 has been associated with clinical trials conducted for the indication of metastatic cancer, where gene therapy strategies were used, and with clinical trials specifically conducted for the indication of metastatic melanoma, where cell therapy strategies were used.

The adverse events myocarditis (10028606), evaluation for heart transplantation (10048769), and cardiac failure chronic aggravated (10066498) have been identified as significant in clinical studies designed for the indication of chronic heart failure.

The adverse event coded as 10016288, febrile neutropenia, has been associated with metastatic cancer in studies utilizing the gene therapy method and with acute lymphoblastic leukemia in studies employing the cell therapy approach.

The adverse event coded as 10019482, hemoglobin (Hb) abnormal, has been associated with metastatic cancer in studies utilizing the gene therapy method and with metastatic melanoma in studies employing the cell therapy approach.

The adverse event coded as 10021789, identified as infection, has been linked to metastatic cancer in studies implementing the gene therapy method and to acute myeloid leukemia in those adopting the cell therapy approach.

The adverse effect of leukaemia with 10024288 is directly associated with clinical trials conducted for the indication of AIDS-related lymphoma.

The adverse event coded as 10025327, lymphopenia, has been associated with metastatic cancer in studies utilizing gene therapy approaches and with metastatic melanoma in clinical trials employing cell therapy methods.

Each ADR included in the study was classified as “serious” or “non-serious” by the principal investigator. As seen in Figure 4.3, the frequency of serious adverse reactions is higher in treatment approaches used in complex indications such as AIDS-related lymphoma and recurrent glioblastoma.

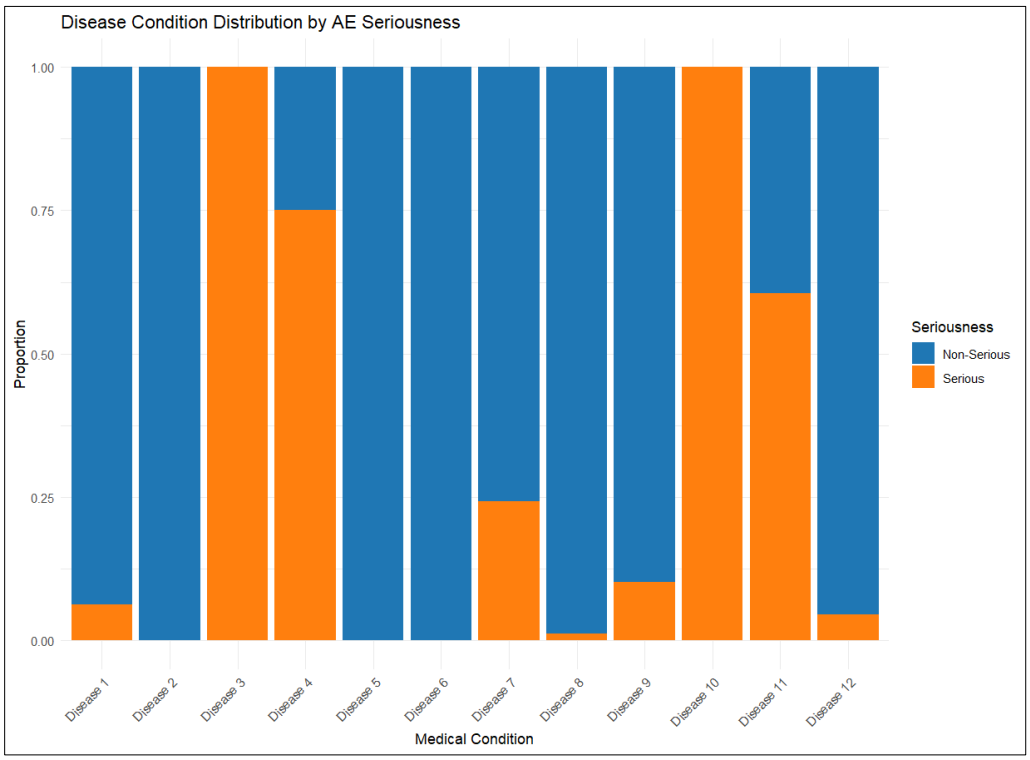


Figure 4.3. The evaluation of disease-related adverse events in terms of seriousness.
Disease 1: Acute Lymphoblastic Leukemia, Disease 2: Advanced Lung Cancer, Disease 3: AIDS-related Lymphoma, Disease 4: Chronic Heart Failure, Disease 5: Chronic Kidney Disease, Disease 6: Hemophilia B, Disease 7: Metastatic Cancer, Disease 8: Metastatic Melanoma, Disease 9: Multiple Myeloma, Disease 10: Recurrent Glioblastoma, Disease 11: Acute Myeloid Leukemia, Disease 12: Transplant rejection

Cancer metastasis accounts for approximately 90% of cancer-related deaths, and despite significant advancements in cancer diagnosis, the treatment of metastatic disease remains a formidable challenge. Platelets, primarily known for their roles in hemostasis and thrombosis, have been increasingly recognized for their pivotal involvement in the metastatic cascade. Circulating tumor cells (CTCs) are bound by activated platelets, which protects them from immune system attacks and prevents them from apoptosis. Additionally, platelets promote the formation of metastatic niches in distant organs by promoting extravasation and improving tumor cell adherence to the endothelium. Platelets induce the epithelial-mesenchymal transition (EMT) in tumor cells by secreting growth factors and cytokines, which increases the cells' capacity for invasion and metastasis. These complex processes highlight how important platelets are to the development and spread of cancer. Thrombocytosis is therefore commonly seen as a predicted outcome in patients with metastatic disease (Zhou et al., 2023).

Elevation in serum ALT and AST levels is a significant indicator of liver dysfunction. In the study NCT01236573, liver dysfunction evidenced by increased serum ALT levels was observed in nearly all cohorts, with the incidence of these events tending to increase at higher dose levels. Among the doses where tumor regression was observed, 5 out of 16 patients experienced grade 3 liver function toxicity, while 3 patients exhibited grade 4 liver toxicity. This highlights the importance of evaluating the adverse event profile during the study to determine the optimal effective dose for patients (Zhang et al., 2015).

Furthermore, considering that a clinical study conducted for a metastatic cancer indication is mentioned, the patient may also need to be evaluated for liver metastasis. GGT, ALT, AST, LDH, and CEA levels were found to be significantly higher in patients with liver metastasis compared to those without, whereas no significant differences were observed in ALP, TP, ALB, GLOB, ALB/GLOB, TBIL, and DBIL levels between patients with and without liver metastasis (Wu et al., 2010).

Acute myelomonocytic and monocytic leukemia often present with renal failure, which is typically caused by a combination of glomerular and tubular dysfunction. This condition is linked to a poor prognosis, and currently, no established treatment guidelines are available. Acute renal failure accompanied by proteinuria, hypokalemia, and hypophosphatemia is distinctly characteristic of patients with AML M4 and M5 subtypes. This specific combination is not observed in other AML subtypes and results from a simultaneous dysfunction of both glomerular and tubular systems (Tapper et al., 2014).

The adverse effect of electrocardiogram QT prolonged with code 10014387 is reported with clinical trials conducted for the indication of multiple myeloma. Cardiovascular events are observed in approximately 7.5% of patients with multiple myeloma, with the type and risk of toxicity varying significantly across studies depending on patient characteristics and the treatments applied. One of the treatment strategies, CAR-T cell therapy, can be associated with direct or indirect cardiac toxicities. CAR-T therapy is linked to an increase in cytokines such as IL-6, IL-2, IL-8, and IF-gamma, which can lead to modifications in the expression and function of calcium and potassium ion channels in myocytes. This increase may be responsible for conditions such as hypotension,

tachycardia, and arrhythmias. These effects arise partly due to the activation of the sympathetic nervous system, leading to cardiac stimulation (El-Cheikh et al., 2023).

Febrile neutropenia is defined as fever and low white blood cell count and is frequently seen in patients with haematological malignancies, including patients undergoing haematopoietic cell transplantation (HCT) or receiving cellular therapies such as chimeric antigen receptor (CAR)-T-cell therapy. In patients with hematologic malignancies, critical effector cells are often impaired and/or reduced in number. Neutropenia (absolute neutrophil count ≤500 cells/μL) is a well-recognized predictor of serious infections and infection-related mortality. The initial approach to a cancer patient presenting with fever and neutropenia involves assessing their risk of infection. The depth and duration of neutropenia remain the most significant risk factors for acute, life-threatening bacterial infections and typically prompt intervention with broad-spectrum empirical antibiotics upon the onset of fever. Patients should be evaluated for sepsis parameters and clinical stability (Stohs et al., 2024).

Low Hb levels negatively impact patients' overall condition and their response to treatment. Cancer-related anemia contributes significantly to diminished quality of life (QoL) by exacerbating symptoms such as fatigue, shortness of breath, and general weakness, making the treatment process more challenging. These findings highlight the importance of monitoring and managing Hb levels as an integral part of cancer care to improve patient outcomes and treatment efficacy. The adverse event coded as 10020949, hypocalcemia, has been associated with metastatic cancer in studies utilizing the gene therapy method and with metastatic melanoma in studies employing the cell therapy approach (VanBelle, 2004).

Anemia is a common complication in cancer and is strongly associated with fatigue and a decline in QoL. Retrospective data analyses from patients with both hematological malignancies and solid tumors indicate that low baseline Hb levels serve as a prognostic marker for poor clinical outcomes, including reduced overall survival (OS) and disease-free survival (DFS) (Sincan et al., 2024).

Certain pre-treatment risk factors in CAR-T cell recipients, such as the type and status of the underlying disease, number of prior chemotherapy lines, and history of allogeneic hematopoietic stem cell transplantation (HSCT), increase infection risk. This risk is further compounded by treatment-related factors, including lymphodepleting conditioning chemotherapy, prolonged cytopenias due to the "off-tumour, on-target" effects of CAR-T cells, and intensified immunosuppression used for managing cytokine release syndrome (CRS) and immune effector cell-associated neurotoxicity syndrome (ICANS) (Stohs et al., 2024).

CAR-T cell therapy recipients experience neutropenia durations comparable to those of autologous HSCT recipients, placing them at a similar risk for infection. Neutropenia is nearly universal due to lymphodepleting conditioning therapy, lasting an average of 10 to 14 days. Consequently, FN within the first 30 days post-CAR-T cell therapy is a common occurrence. The approach to fever and neutropenia in CAR-T cell therapy patients must account for not only infectious complications but also CAR-T cell therapy-related toxicities such as CRS and ICANS. Empirical antipseudomonal beta-lactams remain the standard of care for FN, supported by the recent epidemiology of organisms causing bloodstream infections, so ongoing surveillance of local epidemiology is essential (Stohs et al., 2024).

Disruption of phosphorus balance leading to hypophosphatemia is frequently observed in cancer patients and can lead to severe complications, affecting their overall prognosis. The development of hypophosphatemia is influenced by various factors, such as the severity of the illness, nutritional status, type of cancer, and the treatments being administered. It can occur due to phosphaturic mesenchymal tumors or as a paraneoplastic effect. The clinical manifestations of hypophosphatemia vary based on its duration and severity, impacting multiple organ systems. In addition to other serious consequences, hypophosphatemia can hinder tissue oxygenation and cause issues such as hemolysis, dysfunction of white blood cells and platelets, encephalopathy, seizures, arrhythmias, cardiomyopathy, rhabdomyolysis, and even coma. Several studies have shown that hypophosphatemia serves as a negative prognostic indicator for inpatients, leading to

extended hospital stays, higher mortality rates, and increased postoperative complications (Adhikari et al., 2021).

As life expectancy increases with antiretroviral therapy (ART) in people living with HIV (PLWH), the incidence of age-related cancers, particularly acute myeloid leukemia (AML), is increasing. Although HIV is not directly mutagenic, it may indirectly contribute to the development of AML by suppressing the immune system and disrupting the bone marrow microenvironment. AML usually occurs at a younger age in HIV-positive individuals, and prognosis is closely related to CD4 cell count. Survival is significantly lower in those with CD4 <200 cells/μL. Intensive chemotherapy and allogeneic stem cell transplantation are currently possible in individuals on ART, but the risk of infection is the most important cause of death. A multidisciplinary approach and further research are required in the diagnosis, treatment, and follow-up processes of HIV-associated AML (Forghieri et al., 2020).

According to the framework of this thesis, termination reasons of ATMPs related clinical trials were classified. Related percentages are presented in Figure 4.4

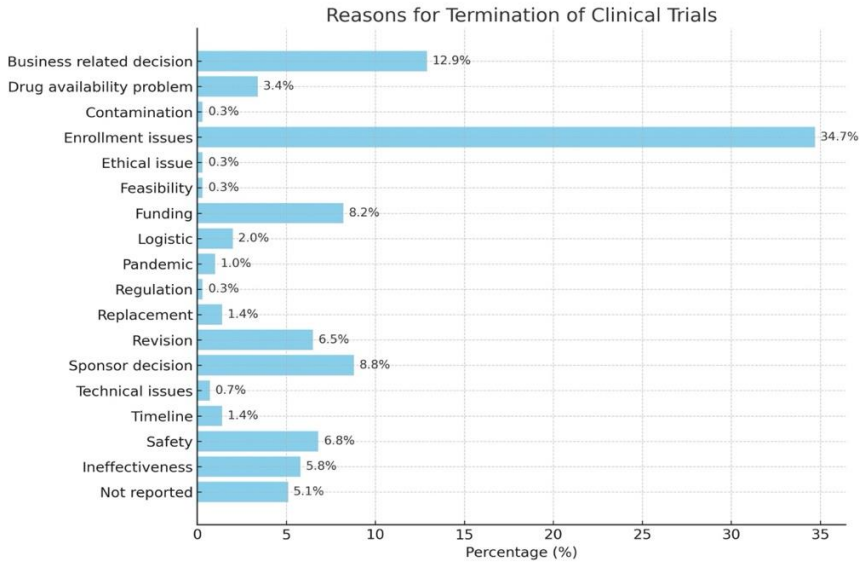


Figure 4.4. Termination reasons of ATMP related clinical trials.
The analysis results indicate that there are various reasons for trial terminations. Enrollment issues are identified as the leading cause of termination, with the highest rate among all reasons. Following closely behind are business-related decisions, funding problems, and sponsor-related decisions. Safety-related terminations, which are the primary focus of this thesis, come immediately after these reasons.

In Figure 4.5., distribution of all ATMP- terminated clinical trials and in Figure 4.6 distribution of terminated clinical trials for safety related terminated studies by phases were presented.

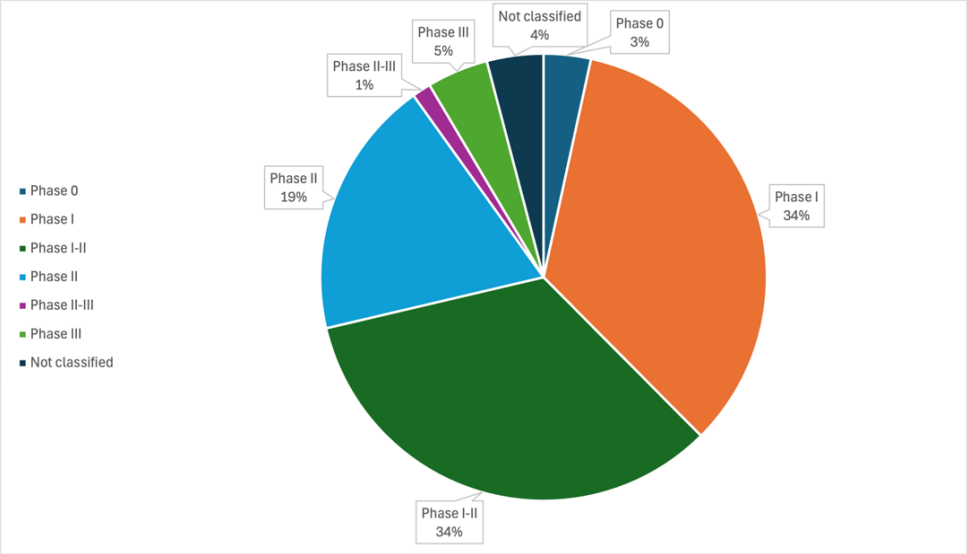


Figure 4.5. Distribution of all ATMP- terminated clinical trials by phases.
Among the 294 identified clinical trials conducted within the scope of ATMPs and terminated for various reasons, 10 (10/294, 3%) were Phase 0, 100 were (100/294, 34%) Phase I-II, 100 (100/294, 34%) were Phase I, 55 (55/294, 19%) were Phase II, 4 (4/294, 1%) were Phase II-III, and 13 (13/294, 5%) were Phase III trials. Additionally, 12 (12/294, 4%) trials were not classified under any specific phase.

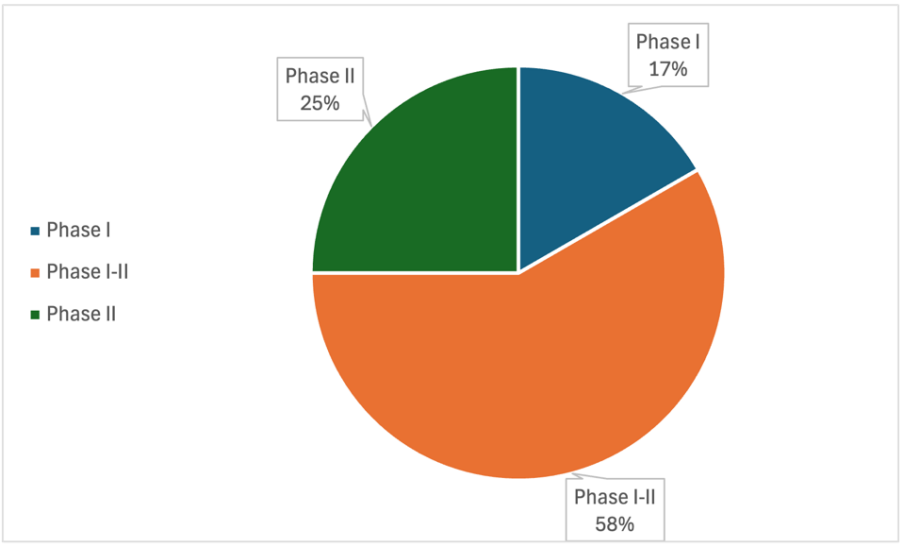


Figure 4.6. Distribution of terminated clinical trials by phases for safety related terminated studies.
Among the 12 clinical trials that were terminated directly due to safety-related concerns, termination occurred in 2 (2/12, 17%) trials at Phase I, 7 (7/12, 58%) trials at Phase I-II, 3 trials at Phase II (3/12, 25%).

As shown in Figure 4.7, safety-related terminated studies are classified according to their methodology. In Figure 4.8, viral vectors have been classified within themselves.

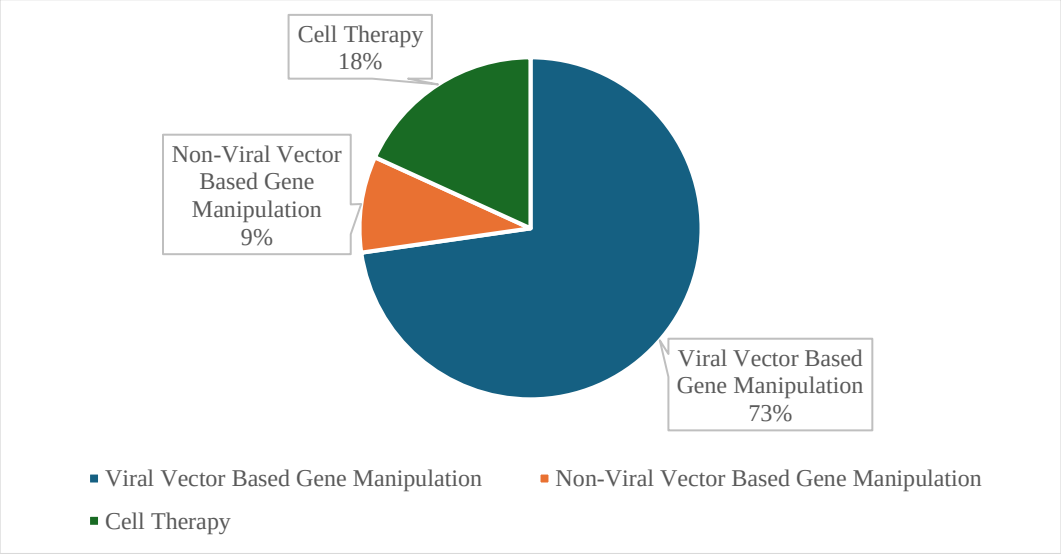


Figure 4.7. Classification of the studies by methodology.
According to the study result, there are 2 (8/12, 73%) viral vector-based gene therapy and 1 (1/12, %9) non-viral based gene therapy, 6 (2/12, 18%) cell therapy.

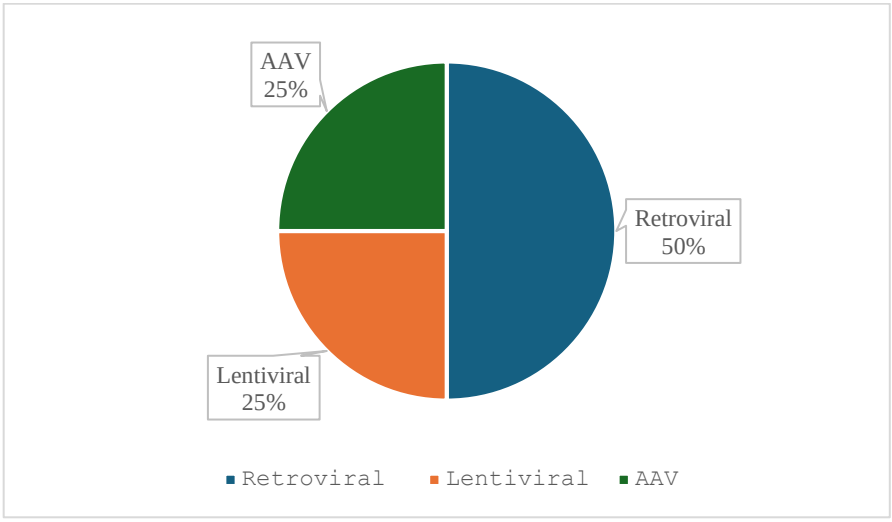


Figure 4.8. Classification of the viral based gene manipulation
Among the viral-based gene manipulation approach, 2 (2/8, 25%) study used AAV, 4 (478, 50%) used retroviral vectors and 2 (2/8, 25%) study used lentiviral vector.

Within the scope of the thesis, studies examining similar clinical outcomes have also been researched from the literature.

Hanna et al. reported that 143 ATMP clinical trials were withdrawn, discontinued or terminated early in their analysis for the period 1999-2015. Between 1999 and 2013, 475 ongoing and incomplete ATMP studies were identified. A termination rate of 23.18% was calculated, similar to the termination rate of conventional products in the clinical trials phase. The probability of stopping clinical trials was 27.35% for cellular therapies, 16.28% for gene therapies and 16.34% for tissue engineered product (TEP) therapies. Oncology had the highest discontinuation rate at 43%, followed by cardiology at 19.2% (Hanna et al., 2016).

Aguilera-Cobos et al. designed a study as a comprehensive umbrella review, analysing data from 71 systematic reviews to identify the main barriers faced by ATMPs in their development, evaluation and transition to clinical practice. Of the 71 systematic reviews, 50 examine safety and efficacy assessment, 13 examine the transition from preclinical to clinical phases, and 8 examine health technology assessment and economic analyses. In total, 13729 clinical trials were analysed. Of the ATMP clinical trials, 11 per cent (approximately 1510/13729) were terminated early. Of these, 41 per cent were stopped due to patient recruitment difficulties and 26 per cent for safety reasons. In addition, 60% of ATMP clinical trials had insufficient numbers of participants to conduct an evaluation. Another 38% of the studies were assessed as low quality due to methodological weaknesses such as small cohorts and single-centre designs (Aguilera-Cobos et al., 2022).

J. Williams et al. conducted a cross-sectional, descriptive study on terminated clinical trials published in the ClinicalTrials.gov results database as of February 2013. Approximately 12% (905/7646) of the trials with published results in the ClinicalTrials.gov database had been terminated. Most of these trials were discontinued for reasons other than accumulated study data (68%, 619/905), with insufficient accrual being the most common reason (57%, 350/619). The remaining trials were terminated

based on study data, including findings related to efficacy or toxicity (21%, 193/905), while 10% (93/905) did not specify a reason for termination (Williams et al., 2015).

In another study, it was aimed to investigate the rates, reasons, and predictors of early termination in oncology clinical trials in the United States. A total of 8687 interventional cancer treatment trials registered on ClinicalTrials.gov between 2007 and 2015, with at least one site in the U.S., were analyzed. The rates, reasons, and predictors of early termination were evaluated using Fisher's exact test, chi-square test, and logistic regression models. The findings revealed that 22.74% of the trials were terminated early, with the most common reasons being accrual issues (34.5%) and "other, multiple reasons, or unknown" (36.9%). Multivariable analysis showed that phase 2 trials, trials funded by academia/foundations, and trials conducted solely in the U.S. were more likely to be terminated early, while trials spanning multiple cancer categories and international trials. were less likely to be terminated. These findings underscore the need to develop strategies to improve patient accrual and enhance the efficiency of oncology clinical trials (Zhang and DuBois, 2023).

Baldi et al. examines the reasons for early termination of cardiovascular clinical trials registered in the ClinicalTrials.gov database between 2006 and 2015, focusing on geographical and temporal patterns. A total of 13729 cardiovascular trials were analyzed, of which 8,900 were classified as closed. Among these, 11% were terminated early, with the most common reason being poor accrual (41%). Intercontinental trials demonstrated higher success in patient recruitment, whereas single-continent trials exhibited higher rates of failed accrual. The findings support the positive impact of globalization on patient enrollment in cardiovascular clinical research. However, early termination results in significant human and economic resource losses, potentially delaying medical progress.

Different study examines the rates of early discontinuation and non-publication of surgical randomized controlled trials. A total of 395 trials registered on ClinicalTrials.gov between 2008-2009 were analysed. 21% of the trials were discontinued early, with poor patient recruitment being the most common reason. Of completed trials, 34% were left unpublished and 66% were published. Funding from the industry decreased the chances of publication. It is challenging to get in touch with researchers, which wastes resources

and raises ethical questions. The study makes suggestions for lowering resource waste and enhancing research transparency (Chapman et al., 2014).

According to the study, 26% of the early-terminated clinical trials were discontinued due to safety-related reasons. The primary factors leading to termination in this category included excessive toxicity, safety concerns, and recommendations from independent Data Safety Monitoring Boards (Baldi et al., 2017).

5. CONCLUSION

The purpose of this thesis is to provide illumination on the relationship between safety concerns and the reasons why clinical studies for ATMPs are terminated. Since Bayesian methods are frequently used to analyse ambiguous data structures by adding previous knowledge, the initial goal was to utilize a Bayesian statistical strategy to investigate this link. However, early in the investigation, some significant restrictions were found that hindered the Bayesian method's successful application.

The dataset's sparseness, which reduced Bayesian inference's robustness, was the main obstacle faced. The analysis was further complicated by the difficulty in establishing meaningful prior distributions due to the lack of adequate prior information. The dataset's weak covariance structure was another important drawback that reduced Bayesian modeling's efficacy. All these elements combined to make the Bayesian technique unsuitable for the required study. The dataset did not match the statistical requirements for a reliable Bayesian analysis.

As a result, a different strategy was chosen, and the study methodology was updated. The Random Forest method, a machine learning approach that works well with complicated datasets with many variables, was finally used to carry out the study. The results were then displayed as figures, which made it possible to visualize the connections between safety issues and termination grounds in an organized and understandable manner. This change in approach guaranteed the validity of the findings while allowing the study to move on in a more exploratory and data-driven manner.

Examining the relationship between adverse effects and the particular ATMPs approach used for certain indications was a one of the other crucial parts of this research. The study was carried out by methodically assessing the data from clinical trials and identifying any adverse events that were recorded. Each phase of a clinical trial serves a unique purpose, with the findings from one phase informing the design of the next.

Ensuring safety is the primary parameter in clinical trials. Concerns about safety emerging among patients can disrupt the anticipated benefit-risk balance of a therapeutic product,

potentially leading to the early termination or suspension of the clinical study, and, in some cases, irreversible outcomes.

It is crucial to point out that the majority of the adverse events found in this study were acute adverse events that manifested during the clinical trial period. These adverse events, which were documented throughout the corresponding research, highlight the short-term or immediate safety concerns with ATMPs.

However, one of the most significant safety-related challenges in the field of ATMPs is the long-term safety profile of these therapies. Due to their innovative and often irreversible nature, ATMPs carry the potential for delayed adverse effects that may occur years after the initial administration. Such long-term consequences, including oncogenic risks, immunogenic responses, or secondary genetic alterations, are among the primary concerns in post-marketing surveillance. Despite the importance of these long-term effects, this study focuses exclusively on data obtained from clinical trials and does not include post-marketing or long-term follow-up data. As a result, long-term adverse effects were not within the scope of the present analysis.

On the other hand, determination of the seriousness of adverse reactions is a decision left to the initiative of the principal investigator. The evaluation of the safety reports of clinical studies conducted in Türkiye as SUSAR is carried out by associating the relevant side effect with the study drug and determining the seriousness criteria. Reasons such as involved or prolonged patient hospitalization, death, significant or permanent sequelae, encountering a life-threatening condition, congenital anomaly may directly provide the serious evaluation of the adverse reaction. In addition, it is also important for the investigator to evaluate the adverse reaction as “serious”. With the Random Forest analysis performed within the scope of the study, it was possible to predict to adverse reactions as “serious” and “non-serious”.

The model made its predictions largely correct in the entire dataset. This shows that the classification performance is generally strong.

In addition, the model's accuracy is above the No Information Rate (NIR = 91%) and is statistically significant (p = 0.02887). This shows that the model makes successful predictions by actually learning the patterns in the data, not just by predicting the majority class.

For class 0 (majority class), the model's sensitivity is 98.96% and the positive predictive value is 95.48%. This shows that the model is quite successful in both capturing and correctly classifying class 0 data. This can be advantageous for areas where early exclusion of negative situations is required, such as clinical or safety.

Since class 1 (minority class) is only 9%, this situation caused the model to learn mostly class 0. With a specificity of 52.63% and NPV of 83.33%, it was weak in identifying examples belonging to class 1. Since the data were obtained from real world evidence, lack of classification of adverse reactions as “serious” is one of the obstacles of this study.

Clinical trials involving ATMPs that were terminated due to safety concerns have identified severe adverse events associated with immunogenicity. When therapeutic cells or gene therapy vectors are recognized as foreign by the immune system, they can trigger excessive immune activation or immunosuppression, leading to various adverse effects. In the clinical studies examined in this research, cytokine release syndrome was the mostly reported adverse event. CRS is a severe inflammatory response characterized by excessive cytokine production, commonly seen in therapies that activate the immune system, such as CAR-T cell therapy. It can lead to hypotension, multiple organ failure, and even death.

Also, encephalopathy is a severe neurotoxicity associated with CRS and neuroinflammation, potentially caused by blood-brain barrier disruption and cytokine effects on the central nervous system.

Immunogenicity remains a major risk factor for the safety of ATMPs. These adverse events highlight the necessity of careful immune response management and the development of more effective immunomodulation strategies in ATMPs approaches.

To further enhance the reliability of the findings, the adverse effects identified through clinical trial data were subjected to an extensive literature search. This search aimed to

distinguish between adverse events that could be directly attributed to the therapeutic intervention and those that might be intrinsic to the natural progression of the disease itself. In other words, scenarios in which the reported adverse effects could be expected outcomes due to the underlying condition were carefully considered.

According to a thorough analysis of the results of research, no previous study has specifically addressed the causes of clinical trial terminations in the field of ATMPs. Even though there are studies assessing clinical trial terminations for various treatment modalities, no study has comprehensively examined ATMP-specific termination reasons, especially as associated to safety issues. This study highlights a significant but previously unidentified element of clinical research in gene and cell therapy, filling a gap in the literature.

Notably, safety concerns, efficacy problems, and recruiting difficulties have been found to be the most frequent causes for termination in deeper studies that look at clinical trial discontinuation across a variety of therapeutic areas. The study's results are consistent with these broad patterns, especially when it comes to the frequency of safety-related trial terminations in gene and cell therapy.

One of the study's main conclusions is that patient recruiting issues are the most common cause of studies using gene and cell therapy approaches terminating. This finding is in line with previous research that has consistently emphasized patient enrolment difficulties as a crucial problem in clinical studies. The pool of eligible individuals is constrained by the very restrictive inclusion criteria that are frequently required due to the complexity and specificity of gene and cell therapy clinical trials.

Several factors contribute to these recruitment challenges. ATMPs related trials often require patients to meet highly specific genetic, immunological, or disease-stage criteria, significantly reducing the number of eligible participants. Additionally, many ATMPs target rare diseases or niche indications, making it inherently difficult to recruit a sufficient number of patients for clinical trials. Another important factor is the high treatment risks associated with these therapies, as potential participants may be hesitant to enrol in trials involving novel treatments with uncertain long-term safety profiles. Furthermore, the

complexity of treatment protocols presents additional challenges, as the administration of ATMPs often involves specialized procedures, such as autologous cell extraction and genetic modification, which require sophisticated clinical infrastructure and a high level of patient commitment.

Given these challenges, recruitment difficulties remain one of the primary barriers to the successful completion of gene and cell therapy clinical trials. Addressing this issue may require adaptive trial designs, increased patient engagement strategies, and expanded multicentre collaborations to enhance recruitment feasibility.

This can be attributed to the systemic nature of metastatic disease, which inherently affects multiple biological systems and predisposes patients to higher rates of complications. In cases where the gene or cell therapy fails to achieve the desired therapeutic effect, disease progression itself may contribute to the observed adverse events. This highlights the interplay between treatment efficacy and safety outcomes, emphasizing the importance of distinguishing treatment-related toxicity from disease-related complications.

This study offers a thorough and methodical examination of the causes of the cessation of clinical trials for gene and cell therapy, with an emphasis on safety issues. With more than 1,000 patients enrolled in trials terminated due to safety concerns, the impact on affected patients was significant even though safety-related discontinuations made up just 6.2% of the trials that were terminated.

The study also shows that difficulties in recruiting patients are the most frequent cause of trial termination, highlighting the necessity of strategic enhancements in trial design and patient enrolment. The results also emphasize how crucial it is to distinguish between adverse reactions caused on by therapy and issues from the disease, especially in cases like metastatic cancer. In the end, this research supports the development of ATMPs in a way that maximizes patient safety and therapeutic efficacy by illuminating these crucial elements and supporting continued efforts to optimize the design, safety assessment, and regulatory oversight of gene and cell therapy clinical trials.

6. REFERENCES

Adedokun OA, Burgess WD. Analysis of Paired Dichotomous Data: A Gentle Introduction to the McNemar Test in SPSS. J Multidiscip Eval. 2012;8:125–31.

Adhikari S, Mamlouk O, Rondon-Berrios H, Workeneh BT. Hypophosphatemia in cancer patients. Clin Kidney J. 2021;14:2304–15.

Aguilera-Cobos L, Rosario-Lozano MP, Ponce-Polo A, Blasco-Amaro JA, Epstein D. Barriers for the evaluation of advanced therapy medicines and their translation to clinical practice: Umbrella review. Health Policy (New York). 2022a;126:1248–55.

Allaire J. RStudio: integrated development environment for R. Boston, MA; 2012.

Baldi I, Lanera C, Berchialla P, Gregori D. Early termination of cardiovascular trials as a consequence of poor accrual: Analysis of ClinicalTrials.gov 2006-2015. BMJ Open. 2017; 7.

Van Belle SJP. What is the value of hemoglobin as a prognostic and predictive factor in cancer? European Journal of Cancer, Supplement. 2004; 2: 11–9.

Benencia F, Sprague L, McGinty J, Pate M, Muccioli M. Dendritic cells the tumor microenvironment and the challenges for an effective antitumor vaccination. J Biomed Biotechnol. 2012; 2012.

Black JE, Kueper JK, Williamson TS. An introduction to machine learning for classification and prediction. Fam Pract. 2023; 40: 200–4.

Blockeel H, Devos L, Frénay B, Nanfack G, Nijssen S. Decision trees: from efficient prediction to responsible AI. Front Artif Intell. 2023; 6: 1124553.

Bui TA, Mei H, Sang R, Ortega DG, Deng W. Advancements and challenges in developing in vivo CAR T cell therapies for cancer treatment. EBioMedicine. 2024; 106: 105266.

Bürkner PC. brms: An R package for Bayesian multilevel models using Stan. J Stat Softw. 2017; 80.

Butt MH, Zaman M, Ahmad A, Khan R, Mallhi TH, Hasan MM, Khan YH, Hafeez S, Massoud EES, Rahman MH, Cavalu S. Appraisal for the Potential of Viral and Nonviral Vectors in Gene Therapy: A Review. Genes (Basel). 2022;13: 1370.

Cappelli F, Castronuovo G, Grimaldi S, Telesca V. Random Forest and Feature Importance Measures for Discriminating the Most Influential Environmental Factors in Predicting Cardiovascular and Respiratory Diseases. International Journal of Environmental Research and Public Health, 2024; 21: 867.

Cesur-Ergün B, Demir-Dora D. Gene therapy in cancer. J Gene Med. 2023; 25.

Chapman SJ, Shelton B, Mahmood H, Fitzgerald JE, Harrison EM, Bhangu A. Discontinuation and non-publication of surgical randomised controlled trials: Observational study. BMJ (Online). 2014; 349.

Cheever MA, Higano CS. PROVENGE (sipuleucel-T) in prostate cancer: The first FDA-approved therapeutic cancer vaccine. Clinical Cancer Research. 2011; 17: 3520–6. PMid:21471425.

Demir Dora D. ALS Hastalığında İleri Tedaviler: Gen ve Hücre Tedavileri. UysalH, ed. Nadir Hastalık Olarak ALS. Ankara: Palme Yayınevi; 2021, p.273-290.

Demir-Dora D. Gen tedavisinde kullanılan viral olmayan taşıyıcı sistemler, Türkiye Klinikleri; 2021, p. 20–6.

Demir-Dora D. Gene Therapy Methods In Thalassemia and Hemogloninopathies.Thalassemia and Hemoglobinopathies Diagnosis, Therapy and Prevention. Canatan D, ed.2024, p. 452–461.

El-Cheikh J, Moukalled N, Malard F, Bazarbachi A, Mohty M. Cardiac toxicities in multiple myeloma: an updated and a deeper look into the effect of different medications and novel therapies. Blood Cancer J. 2023;13.

European Medicines Agency (EMA). Advanced therapy medicinal products: Overview. 2023. [cited 2025 April 21]. Available from: https://www.ema.europa.eu/en/human-regulatory-overview/advanced-therapy-medicinal-products-overview.

Forghieri F, Nasillo V, Bettelli F, Pioli V, Giusti D, Gilioli A, Mussini C, Tagliafico E, Trenti T, Cossarizza A, Maffei R, Barozzi P, Potenza L, Marasca R, Narni F, Luppi M. Acute Myeloid Leukemia in Patients Living with HIV Infection: Several Questions, Fewer Answers. Int J Mol Sci. 2020;21:1081.

Frederico C, Conceição A, Nóbrega C, Mendonça LS. Advanced therapy medicinal products development - from guidelines to medicines in the market. Biotechnol Adv. 2025;83.

Granato AM, Pancisi E, Piccinini C, Stefanelli M, Pignatta S, Soldati V, Carloni S, Fanini F, Arienti C, Bulgarelli J, Tazzari M, Scarpi E, Passardi A, Tauceri F, La Barba G, Maimone G, Baravelli S, de Rosa F, Ridolfi L, Petrini M. Dendritic cell vaccines as cancer treatment: focus on 13 years of manufacturing and quality control experience in advanced therapy medicinal products. Cytotherapy. 2024;26.

Han J, Wang H. Cytokine-overexpressing dendritic cells for cancer immunotherapy. Exp Mol Med. 2024; 56: 2559–68. PMid:39617785.

Hanna E, Rémuzat C, Auquier P, Toumi M. Risk of discontinuation of Advanced Therapy Medicinal Products clinical trials. J Mark Access Health Policy. 2016a;4:10.3402/jmahp.v4.32232.

Iglesias-Lopez C, Agustí A, Obach M, Vallano A. Regulatory Framework for Advanced Therapy Medicinal Products in Europe and United States. Front Pharmacol. 2019;10.

Kandi V, Vadakedath S. Clinical Trials and Clinical Research: A Comprehensive Review. Cureus. 2023;15:e35077.

Lee SY. Using Bayesian statistics in confirmatory clinical trials in the regulatory setting: a tutorial review. BMC Medical Research Methodology 2024 24:1. 2024;24:1–34.

Mantia-Smaldone GM, Chu CS. A review of dendritic cell therapy for cancer: Progress and challenges. BioDrugs. 2013;27:453–68.

Minneci PC, Deans KJ. Clinical trials. Semin Pediatr Surg. 2018;27:332–7.

Muehlemann N, Zhou T, Mukherjee R, Hossain MI, Roychoudhury S, Russek-Cohen E. A Tutorial on Modern Bayesian Methods in Clinical Trials. Ther Innov Regul Sci. 2023; 57: 402.

National Academies of Sciences E and M, Medicine NA of, Sciences NA of, Committee on Human Gene Editing: Scientific M and EC. Somatic Genome Editing. Human Genome Editing: Science, Ethics, and Governance. 2017:1–310.

Pack DW, Hoffman AS, Pun S, Stayton PS. Design and development of polymers for gene delivery. Nat Rev Drug Discov. 2005; 4: 581–93.

Sarker IH. Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Comput Sci. 2021;2:1–21.

Sincan G, Kılıç AF, Sincan S, Erdem F, Sincan G, Kılıç AF, Sincan S, Erdem F. Prevalence and Prognostic Significance of Anemia in Lymphoma. Medical Journal of Bakirkoy. 2024; 20: 42–7.

Sterner RC, Sterner RM. CAR-T cell therapy: current limitations and potential strategies. Blood Cancer J. 2021; 11: 69.

Stohs EJ, Abbas A, Freifeld A. Approach to febrile neutropenia in patients undergoing treatments for hematologic malignancies. Transplant Infectious Disease. 2024; 26.

Tapper EB, Luptakova K, Joyce RM, Tzachanis D. A 78-year-old man with acute myeloid leukemia (AML) and acute renal failure. American Journal of Case Reports. 2014; 15: 364–7.

Team RCore. R Language Definition. 2025. [cited 2025 June 13]. Available from: https://cran.r-project.org/doc/manuals/r-release/R-lang.html.

Vivier E, Rebuffet L, Narni-Mancinelli E, Cornen S, Igarashi RY, Fantin VR. Natural killer cell therapies. Nature. 2024a; 626: 727–36.

Volodina O, Smirnikhina S. The Future of Gene Therapy: A Review of In Vivo and Ex Vivo Delivery Methods for Genome Editing-Based Therapies. Mol Biotechnol. 2024a;67: 425–37.

Waldman AD, Fritz JM, Lenardo MJ. A guide to cancer immunotherapy: from T cell basic science to clinical practice. Nat Rev Immunol. 2020;20:651–68.

Williams RJ, Tse T, DiPiazza K, Zarin DA. Terminated trials in the clinicaltrials.gov results database: Evaluation of availability of primary outcome data and reasons for termination. PLoS One. 2015;10.

Wu XZ, Ma F, Wang XL. Serological diagnostic factors for liver metastasis in patients with colorectal cancer. World J Gastroenterol. 2010;16:4084–8.

Yao B, Zhu L, Jiang Q, Amy Xia H. Safety monitoring in clinical trials. Pharmaceutics. 2013;5:94–106.

Zhang E, DuBois SG. Early Termination of Oncology Clinical Trials in the United States. Cancer Med. 2023;12:5517–25.

Zhang L, Morgan RA, Beane JD, Zheng Z, Dudley ME, Kassim SH, Nahvi A V., Ngo LT, Sherry RM, Phan GQ, Hughes MS, Kammula US, Feldman SA, Toomey MA, Kerkar SP, Restifo NP, Yang JC, Rosenberg SA. Tumor-infiltrating lymphocytes genetically engineered with an inducible gene encoding interleukin-12 for the immunotherapy of metastatic melanoma. Clinical Cancer Research. 2015;21:2278–88.

Zhou L, Zhang Z, Tian Y, Li Z, Liu Z, Zhu S. The critical role of platelet in cancer progression and metastasis. Eur J Med Res. 2023;28. PMid:37770941.

Advanced therapy medicinal products: Overview. 2024. [cited 2024 June 16]. Available from: https://www.ema.europa.eu/en/human-regulatory-overview/advanced-therapy-medicinal-products-overview.

CRAN: Package dplyr. n.d. [cited 2025 March 29]. Available from: https://cran.r-project.org/web/packages/dplyr/index.html.

CRAN: Package tidyr. n.d. [cited 2025 March 29]. Available from: https://cran.r-project.org/web/packages/tidyr/index.html.

The Turkish Medicines and Medical Devices Agency. Klinik Araştırmalarda Güvenlilik Bildirimlerine İlişkin Kılavuz. 2024.

MedDRA-Help to Shape the MedDRA Terminology, [cited 2025 June 23]. Available from: https://www.meddra.org/.

Excel [Computer Software]. Microsoft Corporation. 2021. [cited 2025 March 21]. Available from: https://www.microsoft.com/tr-tr/microsoft-365/excel.

U.S. Food& Drug Administration. Cellular & Gene Therapy Products. 03/20/2023. 2024. [cited 2024 June 16]. Available from: https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products.

		Condition												Total
		1	2	3	4	5	6	7	8	9	10	11	12	
Code	10000081	5	0	0	0	0	0	0	0	1	0	0	3	9
	10000167	0	0	0	0	0	0	8	9	0	0	0	4	21
	10001041	0	0	0	0	0	0	0	0	0	0	2	2	4
	10001497	6	0	0	0	0	0	0	0	0	0	0	0	6
	10001547	0	0	0	0	0	0	2	6	0	0	0	0	8
	10001551	2	0	0	0	0	0	2	0	0	0	0	0	4
	10002034	9	0	0	0	0	0	0	0	3	0	0	0	12
	10002646	0	0	0	0	0	0	1	8	0	0	0	3	12
	10002855	5	0	0	0	0	0	0	0	1	0	0	0	6
	10002948	15	0	0	0	0	0	0	0	0	0	0	0	15
	10003239	3	0	0	0	0	0	0	0	1	0	0	1	5
	10003477	0	0	0	0	0	0	0	5	0	0	0	0	5
	10003549	7	0	0	0	0	0	0	0	0	0	0	0	7
	10003598	0	0	0	0	0	0	0	0	0	0	0	3	3
	10003658	1	0	0	0	0	0	0	0	2	0	0	3	6
	10003988	4	0	0	0	0	0	0	0	0	0	0	3	7
	10003999	0	0	0	0	0	0	0	0	0	0	3	0	3
	10005364	4	0	0	0	0	0	0	0	0	0	0	0	4
	10007882	2	0	0	0	0	0	0	0	0	0	0	0	2
	10008479	0	1	0	0	0	0	0	0	0	0	0	5	6
	10008531	7	0	0	0	0	1	0	5	0	0	0	0	13
	10009119	0	0	0	0	1	0	0	0	0	0	0	0	1
	10010305	18	0	0	0	0	0	1	0	1	0	0	0	20
	10010774	5	0	0	0	0	0	0	0	2	0	0	1	8
	10011224	5	0	0	0	0	0	0	3	0	0	0	1	9
	10011366	0	0	0	0	0	0	0	0	0	0	0	4	4
	10011906	24	0	0	1	0	0	2	0	0	0	1	3	31
	10012218	6	0	0	0	0	0	0	0	0	0	0	0	6
	10012373	6	0	0	0	0	0	0	0	0	0	0	0	6
	10012378	3	0	0	0	0	0	0	0	0	0	0	0	3
	10012727	0	0	0	0	0	0	1	2	0	0	0	7	10
	10012735	17	0	0	0	0	0	0	0	3	0	0	0	20
	10013573	4	0	0	0	0	0	0	1	0	0	0	4	9
	10013781	2	0	0	0	0	0	0	1	0	0	0	0	3

	10013786	2	0	0	0	0	0	0	1	1	0	0	0	4
	10013946	2	0	0	0	0	0	0	0	2	0	0	0	4
	10013963	0	0	0	0	0	0	6	2	0	0	0	6	14
	10013968	3	0	0	0	0	0	0	0	0	0	0	0	3
	10014080	0	0	0	0	0	0	0	0	0	0	0	3	3
	10014387	0	0	0	0	0	0	0	0	7	0	0	0	7
	10014625	14	0	0	0	0	0	0	0	0	0	0	0	14
	10016256	9	0	0	0	0	0	0	4	1	0	0	5	19
	10016288	7	0	0	0	0	0	7	3	4	0	0	0	21
	10016558	0	0	0	0	0	0	1	8	0	0	0	3	12
	10018867	4	0	0	0	0	0	0	0	2	0	0	0	6
	10019063	3	0	0	0	0	0	0	0	0	0	0	0	3
	10019211	10	0	0	0	0	1	0	0	1	0	0	3	15
	10019482	0	0	0	0	0	0	6	9	0	0	0	3	18
	10019483	0	0	0	0	0	0	2	0	0	0	0	0	2
	10020582	0	0	0	0	0	0	1	2	0	0	0	5	8
	10020639	0	0	0	0	0	0	1	0	0	0	0	4	5
	10020670	0	0	0	0	0	0	1	4	0	0	0	0	5
	10020772	6	0	0	0	0	0	0	0	0	0	0	4	10
	10020937	3	0	0	0	0	0	0	0	0	0	0	0	3
	10020943	0	0	0	0	0	0	0	9	0	0	0	0	9
	10020949	0	0	0	0	0	0	2	7	0	0	0	1	10
	10021015	10	0	0	0	0	0	0	0	5	0	0	0	15
	10021018	0	0	0	0	0	0	0	2	0	0	0	6	8
	10021027	8	0	0	0	0	0	0	0	2	0	0	0	10
	10021028	0	0	0	0	0	0	0	0	0	0	0	5	5
	10021036	5	0	0	0	0	0	0	0	0	0	0	0	5
	10021038	0	0	0	0	0	0	1	4	0	0	0	3	8
	10021058	5	0	0	0	0	0	0	0	2	0	0	0	7
	10021059	0	0	0	0	0	0	6	0	0	0	0	0	6
	10021097	3	0	0	0	0	0	4	1	2	0	1	3	14
	10021143	3	0	0	0	0	0	4	1	0	0	1	2	11
	10021789	0	0	0	0	0	0	4	4	0	0	3	9	20
	10022437	8	0	0	0	0	0	0	0	1	0	0	2	11
	10022595	1	0	0	0	0	0	0	0	0	0	0	0	1
	10024264	8	0	0	0	0	0	0	0	0	0	0	0	8

	10024288	0	0	1	0	0	0	0	0	0	0	0	0	1
	10025256	2	0	0	0	0	0	2	0	2	0	0	0	6
	10025327	0	0	0	0	0	0	8	9	0	0	0	5	22
	10027175	3	0	0	0	0	0	0	0	0	0	0	0	3
	10028127	0	0	0	0	0	0	0	0	0	0	3	7	10
	10028154	0	0	0	0	0	0	0	0	0	0	2	0	2
	10028372	8	0	0	0	0	0	0	0	0	0	0	0	8
	10028596	1	0	0	0	0	0	0	0	0	0	0	0	1
	10028606	0	0	0	1	0	0	0	0	0	0	0	0	1
	10028813	16	0	0	0	0	0	1	7	4	0	0	5	33
	10029354	3	0	0	0	0	0	0	0	1	0	0	0	4
	10029366	7	0	0	0	0	0	2	0	0	0	0	0	9
	10029380	0	0	0	0	0	0	0	0	0	0	0	5	5
	10030095	4	0	0	0	0	0	0	0	0	0	0	0	4
	10030124	7	0	0	0	0	0	0	0	0	0	0	0	7
	10033371	3	0	0	0	0	0	0	9	0	0	0	1	13
	10033374	0	0	0	0	0	0	0	0	0	0	1	0	1
	10033425	5	0	0	0	0	0	0	0	0	0	0	7	12
	10034620	4	0	0	0	0	0	0	0	0	0	0	0	4
	10034960	3	0	0	0	0	0	0	0	0	0	0	0	3
	10035528	7	0	0	0	0	0	2	0	2	0	0	0	11
	10035598	3	0	0	0	0	0	2	1	0	0	0	3	9
	10035664	4	0	0	0	0	0	0	0	3	0	0	0	7
	10037087	4	0	0	0	0	0	0	0	0	0	0	0	4
	10037660	6	0	0	0	0	1	0	0	0	0	0	0	7
	10037844	0	0	0	0	0	0	0	3	0	0	0	6	9
	10038435	0	0	0	0	0	0	2	0	0	0	0	0	2
	10038695	0	0	0	0	0	0	0	0	0	0	2	0	2
	10039906	9	0	0	0	0	0	2	0	0	1	0	1	13
	10040047	2	0	0	0	0	0	0	0	0	0	0	0	2
	10040123	0	0	0	0	0	0	6	0	0	0	0	0	6
	10040335	0	0	0	0	0	0	0	9	0	0	0	0	9
	10040369	0	0	0	0	0	0	0	9	0	0	0	0	9
	10040741	3	0	0	0	0	0	0	0	0	0	0	2	5
	10041349	10	0	0	0	0	0	2	0	0	0	0	1	13
	10042661	0	0	0	0	0	0	0	5	0	0	0	0	5

	10043071	2	0	0	0	0	0	0	0	0	0	0	3	5
	10043088	0	0	0	0	0	0	0	0	0	0	4	0	4
	10044565	12	0	0	0	0	0	0	0	0	0	0	0	12
	10046306	0	1	0	0	0	0	0	0	1	0	0	0	2
	10046543	3	0	0	0	0	0	0	0	0	0	0	1	4
	10047700	15	0	0	0	0	0	0	6	0	0	0	6	27
	10047865	0	0	0	0	0	0	0	0	0	0	0	6	6
	10047900	0	0	0	0	0	0	0	4	0	0	0	0	4
	10047942	6	0	0	0	0	0	0	0	0	0	0	0	6
	10048294	4	0	0	0	0	0	0	0	0	0	0	1	5
	10048769	0	0	0	1	0	0	0	0	0	0	0	0	1
	10048959	1	0	0	0	0	0	0	0	0	0	0	0	1
	10048962	5	0	0	0	0	0	0	0	0	0	0	0	5
	10050068	0	0	0	0	0	0	0	0	0	0	0	7	7
	10051055	2	0	0	0	0	0	0	0	1	0	0	0	3
	10051312	0	0	0	0	0	0	0	0	0	0	2	7	9
	10051746	0	0	0	0	0	0	2	0	0	0	0	2	4
	10052015	37	1	0	0	0	0	0	2	0	0	0	0	40
	10052268	2	0	0	0	0	0	0	0	0	0	0	0	2
	10058080	4	0	0	0	0	0	0	0	0	0	0	0	4
	10061428	8	0	0	0	0	0	0	0	0	0	0	0	8
	10061819	0	0	0	0	0	0	0	0	0	0	7	0	7
	10062959	1	0	0	0	0	0	0	0	0	0	0	0	1
	10062975	0	0	0	0	0	0	0	0	0	0	0	3	3
	10066260	0	0	0	0	0	0	0	0	0	0	4	0	4
	10066498	0	0	0	1	0	0	0	0	0	0	0	0	1
	10068429	0	0	0	0	0	0	1	0	0	0	0	0	1
	10069339	2	0	0	0	0	0	0	0	1	0	0	0	3
	10069640	0	0	0	0	0	0	0	0	0	0	0	4	4
	10074170	3	0	0	0	0	0	0	0	0	0	0	0	3
	10090406	0	0	0	0	0	0	0	0	0	0	2	0	2

APPENDIX-1 The analysis result of McNemar which shows the relationship between indication and adverse reaction occurred in safety-related terminated studies

Abdominal pain-10000081, Agitation-10001497, Alanine aminotransferase increased-10001551, Anaemia-10002034, Anxiety-10002855, Aphasia-10002948, Arthralgia-10003239, Asthenia-10003549, Back pain-10003988, Blood bilirubin increased-10005364, Cellulitis-10007882, Chills-10008531, Confusional state-10010305, Constipation-10010774, Cough-10011224, Delirium-10012218, Depressed level of consciousness-10012373, Depression-10012378, Diarrhoea-10012735, Dizziness-10013573, Dry mouth-

10013781, Dry skin-10013786, Dyspepsia-10013946, Dyspnoea-10013968, Encephalopathy-10014625, Fatigue-10016256, Haematuria-10018867, Hallucination-10019063, Headache-10019211, Hypertension-10020772, Hypoaesthesia-10020937, Hypokalaemia-10021015, Hypomagnesaemia-10021027, Hyponatraemia-10021036, Hypophosphataemia-10021058, Insomnia-10022437, International normalised ratio increased-10022595, Lethargy-10024264, Lymphocyte count decreased-10025256, Memory impairment-10027175, Muscular weakness-10028372, Myocardial infarction-10028596, Nausea-10028813, Neutropenia-10029354, Neutrophil count decreased-10029366, Oedema-10030095, Oedema peripheral-10030124, Peripheral sensory neuropathy-10034620, Photophobia-10034960, Platelet count decreased-10035528, Pleural effusion-10035598, Pneumonia-10035664, Pruritus-10037087, Pyrexia-10037660, Seizure-10039906, Sepsis-10040047, Sinus bradycardia-10040741, Somnolence-10041349, Tremor-10044565, Urinary incontinence-10046543, Vomiting-10047700, White blood cell count decreased-10047942, Mental status changes-10048294, Peripheral swelling-10048959, Brain oedema-10048962, Deep vein thrombosis-10051055, Cytokine release syndrome-10052015, Catheter site pain-10052268, Staphylococcal infection-10058080, Decreased appetite-10061428, Neutropenic colitis-10062959, Acute kidney injury-10069339, Candida infection-10074170, Abnormal platelets-10000167, Acute renal failure-10001041, Alanine aminotransferase abnormal-10001547, Anorexia-10002646, Aspartate aminotransferase abnormal-10003477, Fever-10016558, Hypermagnesemia-10020670, Hypoalbuminemia-10020943, Hyponatremia-10021038, Pain-10033371, Serum magnesium abnormal-10040335, Serum phosphate abnormal-10040369, Sweating-10042661, Weight loss-10047900, Atelectasis-10003598, Atrial fibrillation-10003658, Chest pain-10008479, Creatinine decreased-10011366, Diarrhea-10012727, Ecchymosis-10014080, Hyperbilirubinemia-10020582, Hyperglycemia-10020639, Hypokalemia-10021018, Hypomagnesemia-10021028, Mucositis-10028127, Neutrophils-10029380, Pain in extremity-10033425, Rash-10037844, Tachycardia-10043071, Weakness generalized-10047865, Edema limbs-10050068, Neutropenic fever-10051312, Sensory neuropathy-10062975, Hand and foot skin reaction-10069640, Bacteremia-10003999, Multi-organ failure-10028154, Pain abdominal-10033374, Respiratory failure-10038695, Tachypnea-10043088, Disease recurrence-10061819, Acute graft versus host disease-10066260, Diffuse alveolar hemorrhage-10090406, Chronic renal failure-10009119, Death-10011906, Dyspnea-10013963, Electrocardiogram QT prolonged-10014387, Febrile neutropenia-10016288, Hemoglobin abnormal-10019482, Hemoglobin decreased-10019483, Hypophosphatemia-10021059, Hypoxia-10021143, Renal failure-10038435, Serum albumin decreased-10040123, Lower gastrointestinal hemorrhage-10051746, Acute vascular leak syndrome-10068429, Hypocalcemia-10020949, Hypotension-10021097, Infection-10021789, Leukaemia-10024288, Lymphopenia-10025327, Lymphopenia-10028606, Evaluation for heart transplantation-10048769, Cardiac failure chronic aggravated-10066498, Upper respiratory tract infection-10046306

Disease 1: Acute Lymphoblastic Leukemia, Disease 2: Advanced Lung Cancer, Disease 3: AIDS-related Lymphoma, Disease 4: Chronic Heart Failure, Disease 5: Chronic Kidney Disease, Disease 6: Hemophilia B, Disease 7: Metastatic Cancer, Disease 8: Metastatic Melanoma, Disease 9: Multiple Myeloma, Disease 10: Recurrent Glioblastoma, Disease 11: Acute Myeloid Leukemia, Disease 12: Transplant rejection

RESUME

Personal Information

Name	Edanur	**Nationality**	
Surname	Ersöz	**Phone number**	
Date of Birth		**e-mail**	

Educational Information

Institution		**Graduation Date**
Highschool	Mehmet Emin Resulzade Anatolian Highschool	2016
University	Hacettepe University Faculty of Pharmacy	2021

Work Experience

Job Title	**Organisation**	**Duration**
Clinical Research Associate	Novagenix	2021-2022
Operation Specialist	IQVIA	2022-2024
Seniour Pharmacovigilance Specialist	Pleksus	2024-2025
Patient Safety Partner	Roche	2025-

Foreign Languages	**Name of the Exam**	**Score**
English	YÖKDİL	95

List of Publications

- Ersöz E, Demir-Dora D. Unveiling the potential of antisense oligonucleotides: Mechanisms, therapies, and safety insights. Drug Dev Res. 2024 Jun;85(4):e22187. doi: 10.1002/ddr.22187. PMID: 38764172.