Chemotherapy Toxicity Predictor — Feature Importance Analysis

Tier 1: Critical Inputs (Total Importance: 55%)

Dosage (mg/m²) — 30%

The primary determinant of chemotherapy-induced toxicity. Higher dosages directly amplify the cytotoxic impact not only on malignant cells but also on healthy tissues, leading to severe dose-dependent side effects.

Neutropenia (baseline) — 25%

Reflects the patient's existing bone marrow reserve. Pre-treatment neutropenia indicates impaired hematopoiesis, making the patient extremely prone to further myelosuppression and infection risk during chemotherapy.

Tier 2: Key Clinical Context (Total Importance: 30%)

Cancer Type: Leukemia — 10%

Leukemia originates in the bone marrow, inherently reducing hematopoietic capacity. This places such patients at the highest baseline risk for chemotherapy-related toxicity.

Age — 8%

Age represents physiological reserve. Elderly patients often exhibit reduced renal and hepatic clearance, slower marrow recovery, and overall frailty, making them less tolerant to chemotherapy.

Chemo Regimen: FOLFOX — 6%

A multi-agent regimen widely used in colorectal cancer. Known for frequent myelosuppression, neuropathy, and cumulative toxicity risks, particularly with oxaliplatin components.

Cycles Completed — 5%

Indicates cumulative exposure to chemotherapy. Many toxicities, like neuropathy or cardiotoxicity, are cumulative and emerge only after repeated cycles.

Chemo Regimen: CHOP — 4%

This regimen, common in lymphoma treatment, has high myelosuppressive potential and includes doxorubicin, a drug known for cardiotoxicity.

BMI — 3%

Affects drug pharmacokinetics. Extremes of BMI, particularly low values, indicate frailty or sarcopenia, leading to relative overdosing when dosing is based on body surface area.

Tier 3: Secondary Factors (Total Importance: 15%)

Chemo Regimen: Gemcitabine — 3%

A commonly used single-agent drug known to cause dose-limiting neutropenia, making it an important but secondary risk factor.

Sex — 2%

Biological sex can subtly influence metabolism, body composition, and fat distribution, slightly affecting drug clearance rates.

Tumor Stage — 2%

A direct indicator of disease burden. Advanced stages are often associated with poor general health and systemic inflammation, raising the risk of treatment toxicity.

Metastasis Status — 2%

Highly correlated with tumor stage. Confirms systemic disease and adds predictive strength regarding frailty and poor treatment tolerance.

Smoking Status: Former — 1%

Indicates past pulmonary and cardiovascular damage, which can reduce the body's capacity to handle the stress of chemotherapy.

Chemo Regimen: ABVD — 1%

Used in lymphoma treatment; includes bleomycin, a drug known for pulmonary toxicity. Its inclusion adds marginal but meaningful predictive insight.

Smoking Status: Never — 1%

Serves as the baseline comparator. Helps differentiate toxicity tolerance across smoking-related health conditions.

Cancer Type: Colon — 1%

This variable is largely redundant since its predictive impact is already captured by the FOLFOX regimen used predominantly in colon cancer.

Cancer Type: Lung — 1%

Redundant with treatment regimens and smoking status, which already encapsulate the major predictive information.

Cancer Type: Lymphoma — 1%

Predictive value overlaps with chemotherapy regimens such as CHOP and ABVD, which are selected based on this diagnosis.

Tier 4: Non-Predictors (Total Importance: 0%)

These variables show no independent predictive value and are excluded to prevent model noise or data leakage:

Tumor Size: Highly correlated with tumor stage and metastasis; adds no new information.

Genetic Mutation: EGFR: Predicts response to targeted therapies, not general cytotoxic chemotherapy.

Genetic Mutation: KRAS: Relevant to resistance against targeted anti-EGFR drugs, not toxicity.

Genetic Mutation: TP53: Common mutation linked to tumor progression but not toxicity.

Genetic Mutation: BRCA1: Associated with response to PARP inhibitors but not general chemotherapy tolerance.

Tumor Response: An outcome variable measured after treatment — using it would cause data leakage.

Overall Survival (Months): A result of treatment and disease course, not a predictive input.