CROWDPLAT

NASA
Glenn Research Center

CASE STUDY

IN-FLIGHT FOOD SYSTEMS & ASTRONAUT HEALTH

AI/ML predictive modeling for space medicine and crew health risk analysis

10	5	6+
SUBJECT-MATTER EXPERTS	NASA GRC-APPROVED MODELS	PUBLIC DATASETS LEVERAGED

EXECUTIVE SUMMARY

NASA Glenn Research Center (GRC), under the Human Research Program, partnered with CrowdPlat to develop proof-of-concept AI/ML models linking in-flight food system characteristics to astronaut health incidence rates. The work supported the Crew Health and Performance Probabilistic Risk Assessment (CHP-PRA) framework and required models that could translate terrestrial nutrition and health data into spaceflight-relevant risk insights.

BACKGROUND

The central constraint was evidence scarcity: no direct spaceflight dietary dataset existed at the scale needed for conventional modeling. CrowdPlat therefore structured the project around an analogous-data strategy, using public-health datasets, NASA reference data, domain experts, and iterative NASA GRC review to build and validate five condition-specific proof-of-concept models.

PROJECT FOCUS

- Predict incidence rates for at least five medical conditions, including UTI and atrial fibrillation.
- Cover caloric, macronutrient, micronutrient, and hydration variables across mission durations of 0.1 to 3.0 years.
- Deliver documented software, model assumptions, validation outputs, and test results for NASA's modeling team.

CROWDPLAT

NASA
Glenn Research Center

THE CHALLENGE

Predicting Health Risk from Food Systems

No direct spaceflight dietary dataset was available. CrowdPlat bridged the gap with analogous terrestrial datasets and translated nutrition, hydration, and health variables into spaceflight-relevant inputs.

Key objectives

- Model five health conditions using spaceflight-relevant inputs.
- Align public-health variables to mission duration and food-system factors.
- Deliver assumptions, validation outputs, and test results for NASA review.

MODELED CONDITIONS

Five NASA GRC-Approved Models

Condition	Method	Metric
Indigestion	Random Forest	F1: 0.55
Food Allergy	LR / SVM / RF	F1: 0.72
Atrial Fibrillation	LR / RF	F1: 0.60-0.61
Bone Turnover	RF Regressor	RMSE: 0.082
UTI (x2 datasets)	LR / SVM / RF	F1: 0.57-0.62

Classification models report F1 score; the bone-turnover regression model reports RMSE. All models were proof-of-concept outputs built on surrogate public-health datasets and reviewed through NASA GRC iteration.

THE SOLUTION

Interdisciplinary Team and Structured Five-Phase Delivery

Phase 1

Team Assembly

Recruited 10 SMEs, including an in-house data science PM and external experts in epidemiology, food science, biochemistry, and AI/ML.

Phase 2

Data Gathering

Sourced NHANES, KNHANES, NMCD, PLOS, and NASA NEEMO/NLSP public datasets for surrogate modeling.

Phase 3

Prediction Modeling

Built disease-specific Random Forest, Logistic Regression, and SVM models with NASA GRC iteration.

Phase 4

Testing & Calibration

Ran test cases on NASA sample data and refined thresholds through review and feedback cycles.

Phase 5

Documentation

Delivered user guide, model assumptions guide, validation reports, verification results, and test outputs.

CROWDPLAT

NASA
Glenn Research Center

ANALOGOUS DATA STRATEGY

Modeling Under Scarce Evidence

Because no direct spaceflight dietary dataset was available, CrowdPlat aligned terrestrial nutrition and health datasets with spaceflight modeling needs through a controlled surrogate-data workflow.

- Applied BERT-based food-name matching using 768-dimension embeddings to align civilian food records with NASA food logs.
- Calibrated unit conversions to NIH standards and normalized daily aggregate inputs to mirror continuous mission dietary data.
- Filtered outliers using IQR and 95th-percentile methods and corrected severe class imbalance through down-sampling.
- Validated surrogate datasets against NASA-supplied test samples before acceptance.

OUTCOME

Validated POC Models Delivered

- Five condition models: indigestion, food allergy, atrial fibrillation, femur bone turnover, and UTI - reviewed and accepted by NASA GRC.
- Dynamic input coverage: nutrition, hydration, and caloric variables across mission phases of 0.1 to 3.0 years.
- Correlation outputs: maps linking food system characteristics to health incidence rates for each modeled condition.
- Documentation: user guide, model assumptions guide, validation and verification reports, and test results.
- Milestones: all three payment milestones met across the project lifecycle.

CROWDPLAT ADVANTAGE

Why CrowdPlat Was Fit for the Project

1 Specialized Talent

CrowdPlat assembled a 10-member expert team across data science, epidemiology, food science, biochemistry, biomedical engineering, and AI/ML.

2 Structured Delivery

A dedicated in-house data science lead coordinated communication with NASA GRC stakeholders across all five phases, enabling rapid issue resolution.

3 Rigorous Data Science

The team applied food-name embeddings, multi-dataset calibration, outlier filtering, and class-imbalance correction under scarce-evidence conditions.

Space Medicine • AI/ML Modeling • Systems Biology • Analogous Data Strategy
CHP-PRA • Random Forest • NHANES • NASA GRC