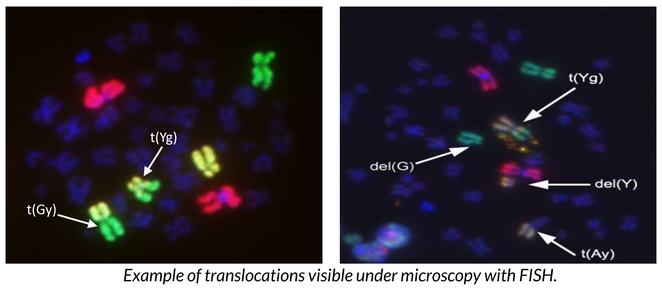
Machine learning-driven dose estimation in astronaut biodosimetry: an alternative to calibration curves in low dose-protracted exposure scenarios

Benjamin Puzantian[1,2], Lindsay A. Beaton-Green[1,2], K. Wu[1], P. Grzegorcyk[1], V. Lai[3], Vincent Cousineau Daoust[4], L. Tomi[4], H. Wu[5], R. C. Wilkins[1,2]

[1]Environmental and Radiation Health Sciences Directorate, Health Canada, Ontario, Canada,[2]Department of Physics, Carleton University, Ontario, Canada, [3]Canadian Nuclear Laboratories, Chalk, Ontario, Canada, [4]Canadian Space Agency, Quebec, Canada, [5]National Aeronautics and Space Administration, Johnson Space Center, Texas, United States of America

BACKGROUND

- Health Canada conducts biodosimetry to quantify doses of ionizing radiation for individuals in several occupational scenarios - including nuclear energy workers and astronauts - using damage to biological materials.
- The Canadian Space Agency (CSA) Operational Space Medicine has set up an Astronaut Biodosimetry Programme in collaboration with Health Canada (HC) in order to better protect and monitor the health of Canadian astronauts.
- Biodosimetry has also been conducted by NASA for their astronauts and by HC for European Space Agency (ESA) astronauts over the past twenty-years [1].
- Chromosomal translocations are sensitive markers that can be detected applying the Fluorescent *in situ* Hybridization (FISH) assay on lymphocyte samples. Traditionally, biological doses from space radiation have been inferred referencing pre-exposure calibration curves using these translocation measurements.



Example of translocations visible under microscopy with FISH.

- Low-dose biodosimetry with calibration curves is limited by using a single endpoint:
  - Excludes confounding factors.
  - May lead to large uncertainty if the frequency of translocations nears the lower limit of detection of the FISH assay.
- The purpose of this work is to **determine whether training a supervised machine learning model with astronaut biodosimetry data could be an alternative method of generating dose estimates.**

MATERIALS AND METHODS

Preflight experiments

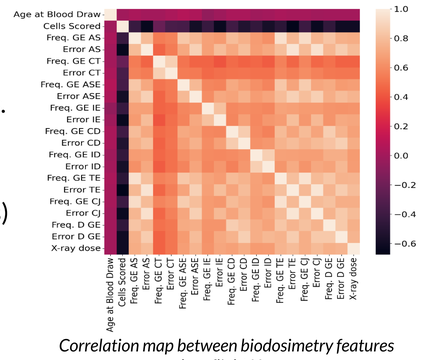
- Biodosimetry data was mined from a comparative study between HC and NASA [1] for astronauts who participated in missions to the International Space Station (ISS) for at least three months.

- Whole blood drawn before flight and irradiated at low LET to generate translocation dose calibration curves:
  - HC samples: cells exposed with cabinet X-ray (2009 – present) and Cs-137 (2007 – 2009).
  - NASA samples: cells exposed with Cs-137.

- Translocations were scored in metaphase lymphocytes (HC and NASA) and prematurely condensed interphase lymphocytes (NASA).

Feature selection

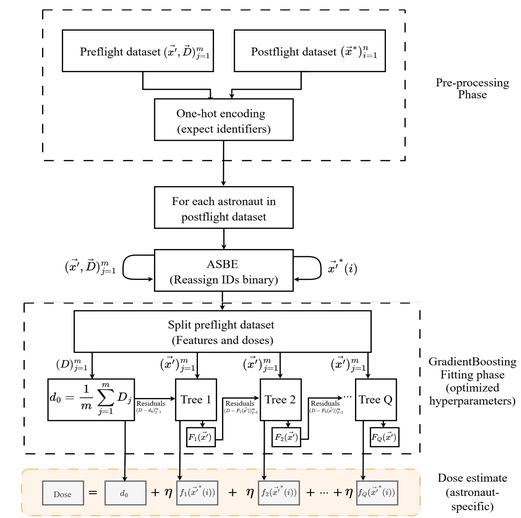
Features engineered from preflight datasets of biodosimetry results.
- 10 features (4 endpoints and 6 confounding factors) were selected from correlation map.



Correlation map between biodosimetry features and preflight X-rays

Supervised machine learning and dose estimation

- Developed astronaut-specific features to consider inter-individual variability in translocation responses.



Schematic representation of the GradientBoosting machine learning approach to astronaut biodosimetry

Dose from missions estimated by:
1. Postflight measured translocation exchanges per lymphocytes compared to preflight calibration curve.
2. Multi-featured approach with GradientBoosting.
3. Effective doses estimated using NASA Space Radiation Cancer Risk Model 2012 (NSCR-2012).

RESULTS

Characteristics of training data

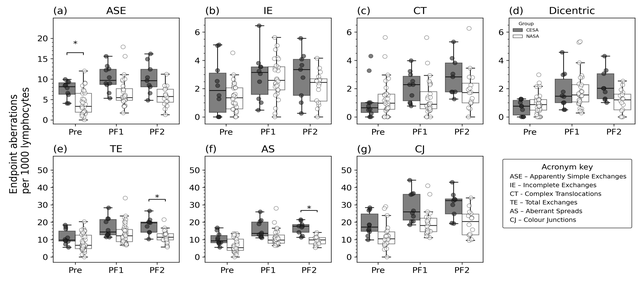
- Biodosimetry results from 25 astronauts were used for machine learning dataset (144 doses between 0-1 Gy).

- CSA-ESA astronaut biodosimetry results denoted CESA.

	CESA (n = 10)	NASA (n = 12 + 3 second flights)
Male: female	82%: 18% (18: 4)	
Age (range) [y]	47 (37 – 56)	46 (35 – 54)
Time in space [d]	180 (± 20)	170 (± 30)
Crew personal dosimeter [mGy]	39 (± 8)	30 (± 6)
Effective dose [mSv]	72 (± 12)	64 (± 15)

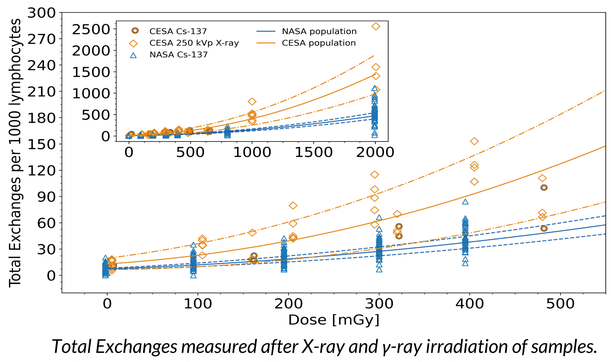
Astronaut demographics and physical dosimetry of ISS missions by agency.

Comparison of endpoints preflight and postflight



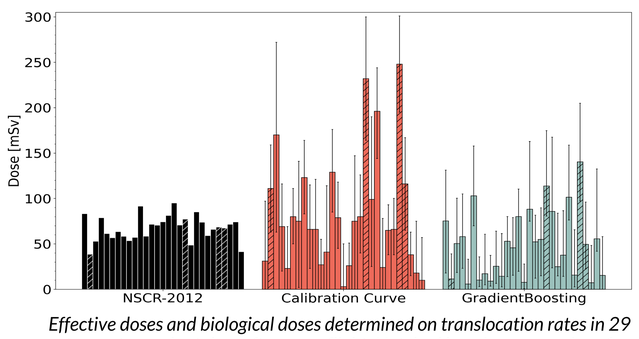
Whisker plots of astronauts' measured endpoint aberrations for 46 astronauts in a cross-comparative study [1]. Translocations persisted in the lymphocyte pool across timepoints. The endpoints are compared to preflight background, showing inter-individual and inter-group variations in rates.

Astronaut dose response calibration curves



Total Exchanges measured after X-ray and γ-ray irradiation of samples. Individual calibration curve fits have been used to infer biological doses. Data taken from the cross-comparative study [1].

Dose estimates



Effective doses and biological doses determined on translocation rates in 29 astronaut samples taken after spaceflight. Hatched bars is a comparison for doses generated with an averaged calibration curve (red). There were fewer large variations between astronauts when using the machine learning.

CONCULSIONS

- A supervised machine learning approach of biodosimetry was successfully developed with the GradientBoosting algorithm, using translocation dose responses scored in lymphocytes of astronauts.
  - The effects of age, sex and inter-laboratory variability were incorporated into the approach.
  - The feasibility of the supervised machine learning analysis was demonstrated and was found to be especially useful when preflight data was insufficient or missing.
  - The machine learning analysis with biodosimetry data brought the dose estimates closer to expected range.

- The supervised machine learning approach was limited by:
  - Lowest dose point of the *ex vivo* blood exposure at 100 mGy.
  - Small sample size resulting in some overfitting.
  - Narrow age range and unequal male/female ratio.

- Future work should focus on optimizing the GradientBoosting model with a broader dataset, containing more granular preflight doses for algorithm training.

- Overall, the machine learning approach provides a possible path to a lower limit of detection in biodosimetry, which could enhance the analysis for similar occupationally exposed cohorts (e.g. nuclear workers) in the low dose-protected settings.

REFERENCES

[1] Puzantian, B. *et al.* Evaluation of chromosome aberrations in astronaut lymphocytes after long-duration missions to the International Space Station. *Life Sciences in Space Research*. In-press. 2026