DISCOVERING SUBGROUPS WITH LARGER WEIGHT GAIN WHEN TAKING DOLUTEGRAVIR

0707

Presenting: Joseph W. Hogan

Jiabei Yang [1], Ann W. Mwangi [2,3], Rami Kantor [4], Issa J. Dahabreh [5], Monicah Nyambura [3], Allison Delong [1], **Joseph W. Hogan** [1], Jon A. Steingrimsson [1]

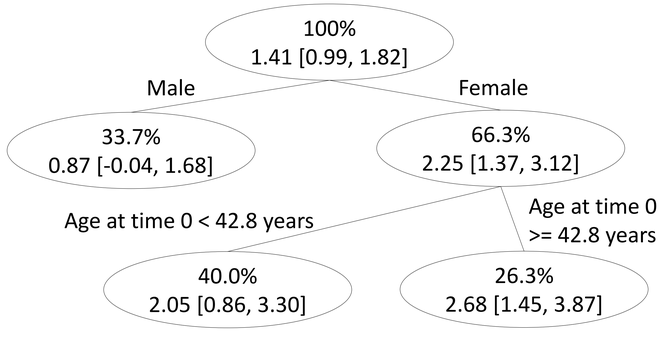
[1] Department of Biostatistics, School of Public Health, Brown University, Providence, RI, US,
[2] Department of Mathematics, Physics & Computing, School of Science & Aerospace Studies, Moi University, Kenya,
[3] Academic Model Providing Access to Healthcare (AMPATH), Kenya,
4 Division of Infectious Diseases, Warren Alpert Medical School, Brown University, Providence, RI, US,
[5] CAUSALab, Department of Epidemiology and Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, US

BACKGROUND

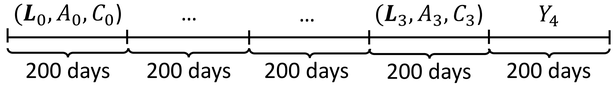
- Clinical trials show that dolutegravir (DTG)-containing antiretroviral therapy (ART) results in larger weight gains than other ARTs.
- Not all people are at the same risk for increased weight gain with DTG-containing ARTs.
- Discovering subgroups for whom DTG is more likely to cause substantial weight gain allows recommending treatment based on a person's risk profile.

METHODS

- We developed Subgroup Discovery for Longitudinal Data (SDLD) algorithm for discovering subgroups with heterogeneous treatment effects using longitudinal data
  - Data-driven tree-based algorithm
  - Combines machine learning and causal inference methods to identify the subgroups with the largest between- group treatment effect difference.
  - Illustration of a tree where in each node
    - 1st row: Percentage of data
    - 2nd row: Estimated treatment effect and 95% CI

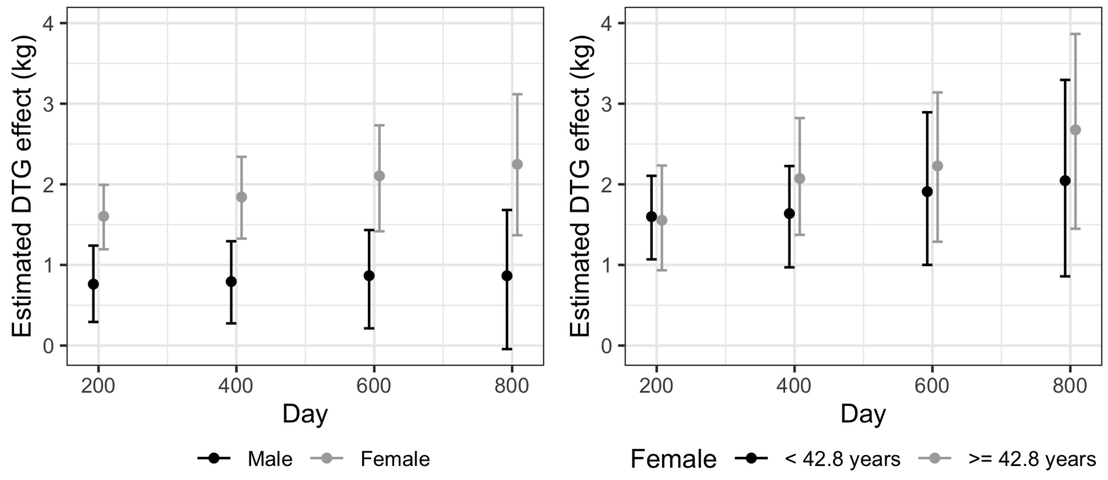


- Use longitudinal data from Academic Model Providing Access to Healthcare (AMPATH) partnership

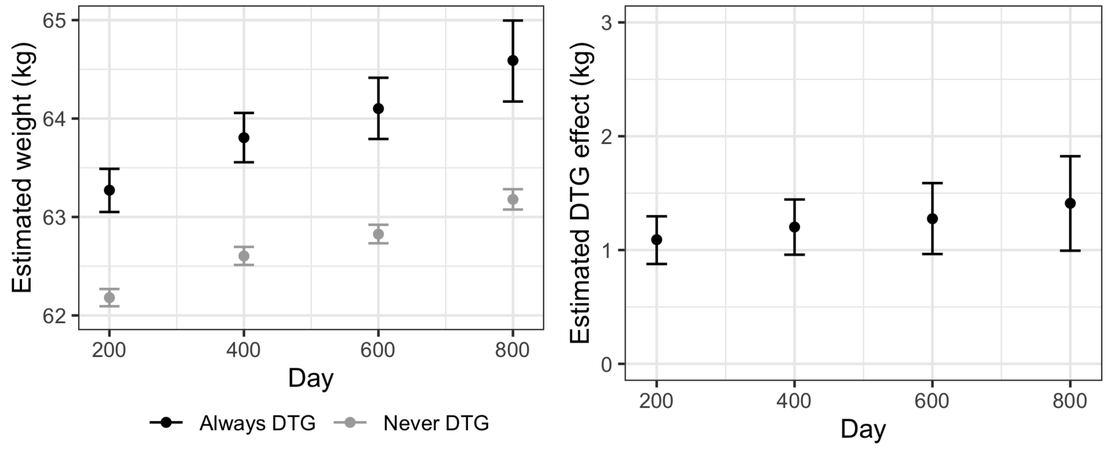


RESULTS

- Data from Academic Model Providing Access to Healthcare (AMPATH) partnership in Eldoret, Kenya
- 84,445 individuals with 1,178,016 clinic visits
- Subgroup Discovery for Longitudinal Data (SDLD) algorithm identified as source of heterogeneity in DTG effects on weight:
  - Left: Gender
  - Right: Among females, whether or not older than or equal to 42.8 years old



- Left: Estimated average weight if always and never on a DTG-containing ART
- Right: Estimated DTG effect on average weight gain comparing always to never on DTG-containing



COVARIATES EXPLORED TO DEFINE SUBGROUPS

Covariates	Baseline Summary
Male, n (%)	28,440 (33.7%)
Age at time 0, years	38.1 ± 10.4
Time on ART at time 0, years	4.0 ± 3.9
Whether enrolled on or after July 1, 2016, n (%)	24,688 (29.2%)
Weight, kg	62.2 ± 12.5
Viral load, n (%)	
Missing	48,309 (57.2%)
$[0, 10^3)$	30,025 (35.6%)
$[10^3, 10^4)$	3,166 (3.7%)
$[10^4, \infty)$	2,945 (3.5%)
Systolic blood pressure, mmHg	118.9 ± 18.6
Diastolic blood pressure, mmHg	73.9 ± 12.0
Height, cm	165.7 ± 9.0
Active TB, n (%)	2,821 (3.3%)
TB treatment, n (%)	1,936 (2.3%)
Married/living with partner, n (%)	47,759 (56.6%)
Covered by NHIF, n (%)	16,485 (19.5%)

- For binary or ordinal covariates, we present the number and proportion of participants falling in the listed category.
- For continuous covariates, we present the mean and standard deviation.
- TB refers to tuberculosis.
- NHIF refers to National Health Insurance Fund.

CONCLUSIONS

- We developed a novel algorithm that combines modern machine learning and causal inference methods to discover subgroups with heterogeneous treatment effects.
- We applied the algorithm to a large EHR database on people living with HIV in western Kenya
- We identified gender as the primary source of heterogeneity for larger weight gain comparing always to never on DTG-containing ARTs.