Supplementary material #1

Trajectory modelling approach

The goal was to model the number of health care contacts per participant per year, and to group participants showing similar patterns of health care visits over time to better capture the intra- and inter-individual heterogeneity of health care utilization.

1. Choosing of a trajectory modelling approach

Considering the health care utilization data used in this study were count data, particularly zero-inflated Poisson (ZIP) data, a group-based trajectory modelling (GBTM) approach (Jones & Nagin, 2007; Jones et al., 2001; Nagin, 1999; Nagin & Odgers, 2010; Nguena Nguefack et al., 2020) was chosen over latent transition analysis (LTA) which mainly handles categorical data, or growth mixture modelling (GMM) which is more suitable for continuous data (Lanza & Cooper, 2016; Muthén, 2004). GBTM is a finite mixture modelling that involves a procedure that gathers individuals into meaningful subgroups that show statistically similar trajectories. Thus, it allows for the possibility of distinct sub-groups within a population and it allows the trajectories to emerge from the data itself (Collins et al., 2014; Jiang, 2015). This offers an alternative to the limitations of using assignment rules based on inherently subjective categorization criteria. The model determines the form and number of groups that best fit the data, and it provides a metric for evaluating the precision of group assignments (Nagin, 1999). GBTM predicts the trajectory of each group and the form of each trajectory, estimates the probability for each individual of group membership (posterior probability), and assigns them to the group for which they have the highest probability to belong (Jones et al., 2001; Nagin, 1999). The SAS® Proc Traj (version 9.4, Cary, NC, USA) was used to perform GBTM. This procedure can be downloaded for free from B. Jones' website (https://www.andrew.cmu.edu/user/bjones/). The full description of the different steps to perform GBTM, as well as the other trajectory modelling approaches can be found elsewhere (Nguena Nguefack et al., 2020).

2. Data preparation

For each participant, the number of health care contacts per year was calculated using administrative data. To work around the convergence problems induced by outliers, the number of health care contacts was set at 50 for any value greater than 50 (applied only for four individuals). The number of health care contacts at three time points (1, 2 and 3 years after completion of CCHS questionnaire) were used to estimate health care trajectories.

3. Choosing the right number of trajectories

Once the GBTM program was applied, the Bayesian information criterion (BIC) was used to select the optimal model (1, 2, 3, 4 or more trajectory groups and different curve possibilities) (Nagin & Odgers, 2010; Schwarz, 1978). The BIC is a measure of the fit of the model that is calculated based on the likelihood of the model and the number of estimated parameters. It favours models that are more parsimonious as compared to the Akaike information criterion (AIC). The model with the lower value (absolute value) of BIC is preferred. The optimal number of trajectories was also chosen to have adequate numbers of participants in each group (at least 5%) (Nagin & Odgers, 2010). The following pages shows the model fit indices used to select the optimal models for the whole study sample and for the various subgroups defined by sex and the GENDER Index tertiles.

4. Trajectories' interpretation

Once the optimal number of trajectories was identified, graphs of observed vs. estimated number of healthcare contacts over time were created and qualitative/clinical descriptors were chosen to describe each trajectory.

Model fit indices tables

Table legend:

[1] 1 = linear (straight line); 2 = linear + quadratic (u-shaped curve/parabola) components

[2] AIC : Akaike information criterion

[3] BIC : Bayesian information criterion

[*] One or more groups in the model had less than 5% of participants

Bold text: Model which best fits the data and respected all criteria (lowest BIC absolute value among trajectory groups that respected the 5% criteria)

Model fit indices for the number of health care utilization trajectories in the whole study sample

Number of trajectories	Trajectory shape [1]	AIC [2] (n=2955)	BIC [3] (n=2955)
1	1	-34331.89	-34343.87
1	2	-34325.65	-34340.63
2	11	-27348.28	-27369.24
2	12	-27345.23	-27369.20
2	21	-27343.57	-27367.53
2	22	-27341.49	-27368.45
3	111	-25943.74	-25973.70
3	112	-25939.86	-25972.81
3	**121**	**-25907.72**	**-25940.67**
3	211	-25943.17	-25976.12
3	122	-25935.23	-25971.17
3	212	-25939.40	-25975.35
3	221	-25937.51	-25973.45
3	222	-25935.29	-25974.23
4	1111*	-25055.74	-25094.68
4	1112*	-25056.73	-25098.67
4	1121*	-25048.88	-25090.82
4	1211*	-25050.08	-25092.02
4	2111*	-25054.82	-25096.76
4	1212*	-25051.08	-25096.01

4	1221*	-25045.90	-25090.83
4	2121*	-25048.00	-25092.93
4	2112*	-25055.81	-25100.75
4	2211*	-25049.80	-25094.74
4	1222*	-25046.59	-25094.52
4	2122*	- 25048.58	-25096.51
4	2212*	-25050.80	-25098.73
4	2221*	- 25045.51	-25093.44
4	2222*	-25046.21	-25097.13

Model fit indices for the number of health care utilization trajectories in males

Number of trajectories	**Trajectory shape [1]**	**AIC [2] (n=1296)**	**BIC [3] (n=1296)**
1	2	-12506.94	-12519.86
2	11	-10297.80	-10315.89
2	12	-10294.35	-10315.01
2	21	-10298.36	-10319.03
2	22	-10295.09	-10318.34
3	111	-9750.51	-9776.35
3	**112**	**-9747.67**	**-9776.09**
3	121	-9749.53	-9777.95
3	211	-9751.31	-9779.72
3	122	-9747.72	-9778.72
3	212	-9748.44	-9779.44
3	221	-9750.42	-9781.42
3	222	-9748.56	-9782.15
4	2211*	-9425.41	-9464.17
4	1222*	-9421.26	-9462.59
4	2122*	-9420.97	-9462.30
4	2212*	-9425.63	-9466.97
4	2221*	-9421.04	-9462.38
4	2222*	-9421.86	-9465.78
5	22222*	-9283.21	-9337.46

Model fit indices for the number of health care utilization trajectories in females

Number of trajectories	**Trajectory shape [1]**	**AIC [2] (n=1659)**	**BIC [3] (n=1659)**
1	2	-22287.79	-22293.20
2	11	-16914.21	-16927.74
2	12	-16914.87	-16931.11
2	21	-16914.87	-16931.11
2	22	-16907.86	-16926.81
3	111	-15867.18	-15888.84
3	**112**	-15867.87	-15892.23
3	121	-15865.80	-15890.16
3	211	-15864.21	-15888.58

3	122	-15866.73	-15893.80
3	212	-15866.73	-15893.80
3	221	-15863.69	-15890.76
3	222	-15864.60	-15894.38
4	1111	-15422.88	-15452.66
4	1112*	-15421.61	-15454.09
4	1121*	-15420.75	-15453.23
4	1211	-15440.30	-15472.79
4	2111	-15445.74	-15478.23
4	2211	-15417.90	-15453.09
4	2121	-15489.51	-15524.70
4	1221	-15417.48	-15452.67
4	**1212**	**-15416.10**	**-15451.29**
4	1222*	-15415.55	-15453.44
4	2122*	-15417.76	-15455.65
4	2212*	-15417.76	-15455.65
4	2221	-15417.92	-15455.82
4	2222*	-15415.93	-15456.54

Model fit indices for the number of health care utilization trajectories in GENDER Index tertile #1

Number of trajectories	Trajectory shape [1]	AIC[2] (n=984)	BIC [3] (n=984)
1	2	-10083.57	-10095.80
2	11	-8300.65	-8317.77
2	12	-8301.25	-8320.81
2	21	-8301.58	-8321.15
2	22	-8302.20	-8324.22
3	111	-7938.86	-7963.32
3	112	-7939.18	-7966.08
3	**121**	**-7936.17**	**-7963.07**
3	211	-7938.86	-7965.77
3	122	-7934.77	-7964.12
3	212	-7939.20	-7968.55
3	221	-7936.74	-7966.09
3	222	-7935.63	-7967.42
4	1111*	-7703.13	-7734.93
4	2111*	-7702.98	-7737.22
4	1211*	-7703.37	-7737.61
4	1121*	-7703.82	-7738.06
4	1112*	-7704.12	-7738.36
4	2211*	-7703.38	-7740.07
4	2121*	-7703.66	-7740.35
4	1212*	-7704.36	-7741.04
4	1221*	-7704.19	-7740.88
4	1222*	-7705.19	-7744.32

4	2122*	-7704.66	-7743.79
4	2212*	-7704.37	-7743.50
4	2221*	-7704.19	-7743.32
4	2222*	-7705.18	-7746.76

Model fit indices for the number of health care utilization trajectories in GENDER Index tertile #2

Number of trajectories	Trajectory shape [1]	AIC [2] (n=986)	BIC [3] (n=986)
1	2	-11545.63	-11557.87
2	11	-9206.04	-9223.16
2	12	-9201.35	-9220.93
2	21	-9201.35	-9220.93
2	22	-9198.26	-9220.28
3	111	-8617.12	-8641.59
3	112	-8618.10	-8645.01
3	121	-8618.11	-8645.02
3	211	-8617.81	-8644.73
3	122	-8619.09	-8648.45
3	212	-8618.79	-8648.15
3	221	-8618.81	-8648.17
3	222	-8619.79	-8651.60
4	1111	-8325.28	-8357.09
4	2111	-8326.02	-8360.28
4	1211	-8318.29	-8352.54
4	1121	-8323.46	-8357.72
4	1112	-8325.34	-8359.60
4	2211	-8319.21	-8355.92
4	2121	-8324.16	-8360.86
4	2112	-8326.09	-8362.79
4	1221	-8312.73	-8349.43
4	1212	-8318.35	-8355.05
4	1122	-8324.26	-8360.96
4	1222	-8313.65	-8352.80
4	2122	-8324.96	-8364.11
4	2212	-8319.28	-8358.42
4	2221	-8313.66	-8352.81
4	2222	-8314.59	-8356.18
5	11111*	-8154.15	-8195.75
5	11222	-8139.19	-8185.68
5	12212	-8151.35	-8197.84
5	12221*	-8136.98	-8183.47
5	21221*	-8143.25	-8189.74
5	22121*	-8143.25	-8189.74
5	22211*	-8323.21	-8369.70
5	21122	-8141.07	-8187.56

5	22112	-8153.23	-8199.72
5	22111*	-8155.14	-8199.19
5	21211*	-8330.02	-8374.06
5	21121*	-8146.94	-8190.98
5	21112*	-8155.00	-8199.05
5	12112	-8152.39	-8196.43
5	11212*	-8152.64	-8196.68
5	**11122**	**-8140.22**	**-8184.26**
5	12211*	-8322.29	-8366.33
5	11221*	-8142.29	-8186.34
5	21111*	-8156.61	-8198.21
5	12111*	-8154.21	-8195.80
5	11211*	-8329.28	-8370.88
5	11121*	-8146.08	-8187.68
5	11112*	-8154.15	-8195.75
5	12222	-8136.02	-8184.96
5	21222	-8140.04	-8188.98
5	22122*	-8142.45	-8191.39
5	22212	-8152.27	-8201.21
5	22221*	-8137.94	-8186.88
5	22222	-8136.98	-8188.36

Model fit indices for the number of health care utilization trajectories in GENDER Index tertile #3

Number of trajectories	**Trajectory shape [1]**	**AIC [2] (n=985)**	**BIC [3] (n=985)**
1	2	-13215.39	-13222.73
2	11	-9830.89	-9843.13
2	12	-9817.74	-9832.41
2	21	-9831.27	-9845.95
2	22	-9818.50	-9835.63
3	111	-9269.62	-9289.19
3	112	-9261.57	-9283.59
3	121	-9264.52	-9286.54
3	211	-9270.19	-9292.21
3	122	-9258.96	-9283.42
3	212	-9262.08	-9286.55
3	221	-9265.52	-9289.98
3	222	-9259.93	-9286.84
4	1111*	-8981.82	-9008.73
4	2111*	-8981.22	-9010.58
4	1211*	-8981.22	-9010.58
4	1121*	-8968.20	-8997.56
4	**1112**	**-8980.34**	**-9009.69**
4	2211*	-8981.73	-9013.53
4	2121*	-8967.54	-8999.34

4	2112*	-8969.19	-9000.99
4	1221*	-8969.13	-9000.93
4	1212*	-8980.36	-9012.16
4	1122*	-8969.19	-9000.99
4	1222*	-8970.12	-9004.36
4	2122*	-8968.53	-9002.78
4	2212*	-8980.24	-9014.49
4	2221*	-8968.54	-9002.79
4	2222*	-8969.53	-9006.22

Supplementary material references

Collins, J. E., Katz, J. N., Dervan, E. E., & Losina, E. (2014). Trajectories and Risk Profiles of Pain in Persons with Radiographic, Symptomatic Knee Osteoarthritis: Data from the Osteoarthritis Initiative. *Osteoarthritis Cartilage, 22*(5), 622–630. https://doi.org/doi:10.1016/j.joca.2014.03.009.

Jiang, J. (2015). *Group-based trajectory modeling for longitudinal data of healthcare financial charges in patients with inflammatory bowel disease* University of Pittsburgh]. Graduate School of Public Health. http://d-scholarship.pitt.edu/25293/1/JJIANG_MS_ETD_061715V6-1.pdf

Jones, B. L., & Nagin, D. S. (2007). Advances in Group-Based Trajectory Modeling and an SAS Procedure for Estimating Them. *Sociological Methods & Research, 35*(4), 542-571. https://doi.org/10.1177/0049124106292364

Jones, B. L., Nagin, D. S., & Roeder, K. (2001). A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. *Sociological Methods&Reserch, 29*(3), 374-393.

Lanza, S. T., & Cooper, B. R. (2016). Latent Class Analysis for Developmental Research. *The Society for Research in Child Development, 10*(1), 59–64. https://doi.org/10.1111/cdep.12163

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 345-368). CA: Sage. https://doi.org/http://dx.doi.org/10.4135/9781412986311.n19

Nagin, D. S. (1999). Analyzing Developmental Trajectories: A Semiparametric, Group-Based Approach. *Psychological Methods, 4*(2), 139-157.

Nagin, D. S., & Odgers, C. (2010, 4 January 2010). Group-Based Trajectory Modeling in Clinical Research. *Annu Rev Clin Psychol, 6*, 109-138. https://doi.org/10.1146/annurev.clinpsy.121208.131413

Nguena Nguefack, H. L., Pagé, M. G., Katz, J., Choinière, M., Vanasse, A., Dorais, M., Samb, O. M., & Lacasse, A. (2020). Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches. *Clin Epidemiol, 12*, 1205.

Schwarz, G. (1978). Estimating the Dimension of a Model. *The annals of statistics, 6*(2), 461-464.