Exploring the dynamic transitions of polysubstance use patterns among Canadian youth using Latent Markov Models on COMPASS data

Yang Yang,[a,*] Zahid A. Butt,[a] Scott T. Leatherdale,[a] Plinio P. Morita,[a] Alexander Wong,[a] Laura Rosella,[b] and Helen H. Chen[a]

[a]University of Waterloo, Canada
[b]University of Toronto, Canada

The Lancet Regional Health - Americas 2022;16: 100389

Published online 12 November 2022

https://doi.org/10.1016/j.lana.2022.100389

Summary

Background Understanding what factors lead to youth polysubstance use (PSU) patterns and how the transitions between use patterns can inform the design and implementation of PSU prevention programs. We explore the dynamics of PSU patterns from a large cohort of Canadian secondary school students using machine learning techniques.

Methods We employed a multivariate latent Markov model (LMM) on COMPASS data, with a linked sample (N = 8824) of three-annual waves, Wave I (W_I, 2016–17, as baseline), Wave II (W_{II}, 2017–18), and Wave III (W_{III}, 2018–19). Substance use indicators, i.e., cigarette, e-cigarette, alcohol and marijuana, were self-reported and were categorized into never/occasional/current use.

Outcomes Four distinct use patterns were identified: no-use (S1), single-use of alcohol (S2), dual-use of e-cigarettes and alcohol (S3), and multi-use (S4). S1 had the highest prevalence (60.5%) at W_I, however, S3 became the prominent use pattern (32.5%) by W_{III}. Most students remained in the same subgroup over time, particularly S4 had the highest transition probability (0.87) across the three-wave. With time, those who transitioned typically moved towards a higher use pattern, with the most and least likely transition occurring S2→S3 (0.45) and S3→S2 (<0.01), respectively. Among all covariates being examined, truancy, being measured by the # of classes skipped, significantly affected transition probabilities from any low→high (e.g., $OR_{S2\rightarrow S4} = 2.41$, 95% CI [2.11, 2.72], $p < 0.00001$) and high→low (e.g., $OR_{S3\rightarrow S1} = 0.38$, 95% CI [0.33, 0.44], $p < 0.00001$) use directions over time. Older students, blacks (vs. whites), and breakfast eaters were less likely to transition from low→high use direction. Students with more weekly allowance, with more friends that smoked, longer sedentary time, and attended attended school unsupportive to resist or quit drug/alcohol were more likely to transition from low→high use direction. Except for truancy, all other covariates had inconsistent effects on the transition probabilities from the high→low use direction.

Interpretation This is the first study to ascertain the dynamics of use patterns and factors in youth PSU utilizing LMM with population-based longitudinal health surveys, providing evidence in developing programs to prevent youth PSU.

Funding The Applied Health Sciences scholarship; the Microsoft AI for Good grant; the Canadian Institutes of Health Research, Health Canada, the Canadian Centre on Substance Abuse, the SickKids Foundation, the Ministère de la Santé et des Services sociaux of the province of Québec.

Keywords: Polysubstance use; Use pattern; Dynamic transition; Risk factor; Canadian adolescents; Latent Markov model

Introduction

Polysubstance use (PSU) refers to using multiple addictive substances simultaneously or within a specified period.[5] According to recent evidence, like in many other countries, youth PSU is an ongoing problem in Canada.[6,7] Proceeding work from the COMPASS study (https://uwaterloo.ca/compass-system/), a large prospective cohort study of a convenience sample of Canadian secondary school students, found that in the 2017–18 school year, 18% reported dual-use or multi-

Abbreviations: Cq, COMPASS questionnaire; LASSO, least absolute shrinkage and selection operator; LCA, latent class analysis; LMM, latent Markov model; LTA, latent transition analysis; MI, multiple imputation; ML, machine learning; PSU, polysubstance use

*Corresponding author.
E-mail address: y24yang@uwaterloo.ca (Y. Yang).

Research in context

Evidence before this study

We searched PubMed and Google Scholar using various combinations of the search terms (((substance use) AND (youth OR adolescent*)) AND (pattern* OR trajector*)) AND ((dynamic OR transition)) with no language restrictions for all studies from these databases up to November 27, 2021. The size and scope of existing evidence vary significantly. Previous studies have identified use patterns, associated risk factors, and longitudinal trajectories of substance use in adolescence. Latent Class Analysis (LCA) and Latent Transition Analysis (LTA) are commonly used for identifying use patterns and dynamics based on cross-sectional and longitudinal evidence.[1,2] Substance use indicators often include tobacco, alcohol, and marijuana. Current evidence suggests typical three-use patterns, e.g., no/low use, alcohol use, poly-use, or four patterns, e.g., low-use, one- or dual-use, moderate multi-use, high multi-use.[3] The evidence also reveals that youth are most likely to remain in the same use pattern subgroup and typically transition to a higher use group as they grow older.[4] Limited evidence was found in the literature on the factors that impact dynamic transitions in use patterns.

Added-value of this study

To our knowledge, this population-based study is the first to apply dynamic modelling techniques, Latent Markov Model (LMM), to examine the transition of PSU patterns over time among youth, accounting for student-level characteristics and school-level (environmental) factors simultaneously. The four distinct PSU patterns among Canadian adolescents were no-use (S1), single-use of alcohol (S2), dual-use of e-cigarettes and alcohol (S3), and multi-use (S4). Although S1 had the highest prevalence (60.5%) at Wave I, with time, S3 became the prominent use pattern (32.5%) by Wave III. The marginal distribution of S1 constantly decreased across the three-wave (0.60→0.39→0.25), and that of S3 (0.14→0.25→0.33) and S4 (0.05→0.12→0.20) steadily increased over time, indicating a general tendency towards increasing use for dual and multiple substances. Although S4 had been the minor use pattern across the three-wave, it is alarming that the prevalence increased by 4.5 times over time, and by Wave III, its prevalence became very close to S2 and S1. Regarding the dynamics, S4 was the most stable use pattern, followed by the S3 and S1 subgroups. Among these four patterns, S2 was the least stable pattern. Examining factors and estimates that lead to the dynamics of use patterns over time reveals that the factors were multifaceted and complex across the four use patterns across the three-wave. Among all covariates being examined, truancy, being measured by the # of classes skipped, significantly affected bi-directional transition probabilities over time. With the inclusion of e-cigarettes as an emerging substance for modelling the dynamics of use patterns, we verified that use patterns change with time, and so does the evidence in use patterns. It is recommended that these models be applied to any content area with similar longitudinal data to address more scientific research questions that include complicated transitions with latent processes over time, such as mental health or behaviour change, that can better inform the management and treatment of addiction and other health issues.

Implications of available evidence

The dynamics of PSU patterns in adolescence can inform school health policymakers and intervention experts on dealing with relevant health threats at this developmental stage and throughout the process. Youth residing in the low or intermediate-use pattern groups were most likely to transition to a higher-level use group. An early detection-prevention approach could be initiated with a more effective strategy for at-risk students. Available evidence indicates that the diverse associations between PSU and multifaceted modifiable factors should be considered when designing and implementing interventions targeting multiple youth behaviours.

use of substances, and 16% used single substance in the past 30 days.[7] Studies of this kind indicate that approximately 60% of high school students have not used substances for the past 5 years.[6,7] Although the number of non-user has remained stable, the multi-use of substances cohort is on the rise, possibly due to the emerging trend of e-cigarette use[6] and the high prevalence of youth cannabis consumption across legalisation that occurred in Canada in 2018.[8–10] Studies have shown no significant increase in ever-use of cannabis among youth post-legalisation as of yet,[10] steadily increasing from 30.5% in 2016–17 to 32.4% in 2018–19.[8]

Previous studies of PSU have identified common use patterns among youth as no or low use, alcohol use (i.e., alcohol only or predominantly alcohol use), and multi-use.[11] Most studies focus on tobacco, alcohol, and marijuana consumption due to their high prevalence among youth. For example, a study of Canadian adolescents aged 12–18 in Victoria, British Columbia, examined the past year's substance use and identified three use patterns: low/no-use, dual-use of marijuana + alcohol, and multi-use of cigarettes + alcohol + marijuana + other illicit drugs.[12] E-cigarettes have not been considered in many of these studies due to their newness. However, their popularity has surged among youth in recent years and may contribute to a rise in youth PSU.[6,7,13] Recent research identifies classes of use that involve dual and multi-use e-cigarettes with other substances, indicating the importance of considering these devices when examining multiple substances use.[7]

Age, sex, and ethnicity are the primary individual-level risk factors impacting adolescent polysubstance users in the literature. With age, the older the students, the higher their risk of using multiple substances.[6,7] Additionally, early substance use is a risk factor for becoming polysubstance users in the future.[14] While evidence concerning age as a risk factor is apparent, sex and ethnicity in youth polysubstance use are inconsistent. Other individual-level factors that may influence the risk of youth substance use have also been explored, including mental illness,[15,16] sedentary lifestyle,[17] eating habits,[18] social connectedness,[19–21] and family/peer/school influence.[11,22] Parental drinking and peer effect have both been identified to correlate with multi-use positively.[11,22] Population-level factors such as living in a non-urban setting are associated with multi-use involving predominantly tobacco use.[23] Among studies that have considered socioeconomic status (SES), their results are inconsistent.[7,23,24]

The COMPASS study is based on school settings, collecting hierarchical (student-level and school-level) health data via anonymous COMPASS Questionnaires (Cq).[25,26] The COMPASS system facilitates collecting, translating, and exchanging student- and school-level data from a large sample of secondary school students and their participating schools across several provinces in Canada each school year.[25,26] From a methodological perspective, the existing literature using COMPASS data primarily applied latent class analysis or latent profile analysis to identify single substance use patterns. To date, none of the studies that used COMPASS data examined the *transition* of PSU patterns among youth over time[6] nor explored *risk factors* affecting the dynamics based on student characteristics and school environment perspectives simultaneously. We aim to explore the dynamic transitions of PSU patterns across time on COMPASS data and address the gap that limited evidence exists to examine the factors that impact the dynamics of PSU patterns among youth. Within the scope of the present study, PSU refers to the use of cigarettes, e-cigarettes, alcohol, and marijuana.

Methods

Study design and participants

This retrospective cohort study used COMPASS data, a de-identified health survey collecting student- and school-level information from a large convenience sample of Canadian secondary schools and students each year from 2012 to 2013.[25,26] Parental/guardian consent is required for participation, employing the active-information passive-consent protocols,[27] with active assent from participating students. Participants complete the cover page of the Cq to generate a unique code that allows researchers to link data collected from the same student across multiple years of participation.[25] Anonymisation using unique self-generated codes is perhaps the principal strategy for ensuring COMPASS data remains confidential throughout the remainder of its life cycle. In this study, the three-year linked samples of COMPASS data include Wave I (W_I, 2016–17), Wave II (W_{II}, 2017–18), and Wave III (W_{III}, 2018–19) collected in the school years before the onset of the COVID-19 pandemic. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline. The COMPASS study has received ethics clearance from the University of Waterloo Office of Research Ethics (ORE 30118).

Dataset and data preprocessing

The longitudinal dataset being analysed contains data from 9307 Canadian students from grades 9 to 10 at W_I (including younger students at secondary I through V in Quebec), followed through three-consecutive-year. The participating students were from 76 secondary schools located in Ontario, Quebec, British Columbia, and Alberta. The analyses were restricted to 8824 students with regular patterns in their grade levels, referring to the advancement of students from one grade to another each school year. The COMPASS study uses grades to be relevant to school planners who make plans based on grades. Thus, a student's grade level is a proxy of their age throughout this study. The Cq data contains demographic and personal information and student responses to multiple-choice questions regarding their behaviour and perspectives on health and wellness topics. Community-level data, i.e., school-level socioeconomic status, urbanity, and built environment (BE), are linked to each participating school. Several data preprocessing steps were taken to prepare the data for analysis, including data cleaning, linking, merging, and missing data analysis. Multiple imputations (MI) for missing values were performed with detailed descriptions in Supplementary Materials.

Substance use indicators

Substance use indicators, including cigarette, e-cigarette, alcohol, and marijuana, were assessed using the COMPASS Cq. Cq posed two questions for cigarette and e-cigarette smoking to determine the incidence and frequency of these substances. Alcohol and marijuana consumption frequency was measured using substance-specific measures within the Cq.

Statistical analysis

The LMM was employed to test hypotheses that subgroups of youth tend to differ in their PSU patterns over time. LTA is considered an LMM, and there is no fundamental difference between these two modelling techniques. LMM is the foundation of LTA, combining multivariate (multiple indicators) categorical latent variable models and Markov chain models.[28] Before model fitting, we applied the least absolute shrinkage and selection operator (LASSO) method[29] to select a subset of covariates with detailed descriptions in Supplementary Materials. Starting with the basic version of the LMM without covariates, we added all covariates selected from the LASSO regularization into model-fitting. The initial full model was used to obtain the optimal number of latent statuses (classes). To fine-tune the full model and obtain the best-fitted model, we considered several models nested in the full model by removing covariates that were inconsistently significant in their effects on the initial and transition probabilities one by one and by pairs from subsequent model fitting. The model selection was based on the Bayesian Information Criteria (BIC) value. The goodness-of-fit was measured to evaluate the quality of the fitted models as an additional assessment to BIC. To assess the significance of predictors on the effect of subgroup membership for the initial and transition probabilities, Wald test statistics (t-test) was performed based on the parameter estimates and standard errors.

Statistical analysis was performed using the R language, open-source software for statistical computing and graphics.[30] In particular, the LMest package[31] for generalized LMMs was utilized. RStudio Server 1.4 was set up on Ubuntu 18.04 with a 64 GiB RAM virtual machine running on Microsoft Azure.

Role of the funding source
The funders of this study had no role in study design, data collection, data analysis, data interpretation, or in the writing, review, or approval of the paper. The first, second, third, and final authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Descriptive statistics
At baseline, sex, grade, race/ethnicity, province, urbanity, and household income are time-invariant factors representing characteristics of participating students. Table 1 demonstrates the characteristics of the linked

A. Baseline descriptive at Wave I (2016–17)

Characteristic	Category	TOTAL	
		N = 9307	100 (%)
Sex	Female	4984	53.6
	Male	4272	45.9
	Missing	51	0.5
Grade (at baseline)	7[a]	753	8.1
	8[a]	669	7.2
	9	4594	49.4
	10	3107	33.4
	11	152	1.6
	12	14	0.1
	Missing	18	0.2
Race/Ethnicity	White	6873	73.8
	Black	279	3.0
	Asian	633	6.8
	Latin American	197	2.1
	Other	1282	13.8
	Missing	43	0.5
Province	Alberta (AB)	444	4.8
	British Columbia (BC)	439	4.7
	Ontario (ON)	6255	67.2
	Quebec (QC)	2169	23.3
Urbanity[b]	Rural	26	0.3
	Small urban	2911	31.3
	Medium urban	1339	14.4
	Large urban	5031	54.1
Household income	\$25K–\$50K	1462	15.7
	\$50K–\$75K	4356	46.8
	\$75K–\$100K	3076	33.1
	>\$100K	413	4.4

(Table 1 continues on next page)

(Continued from previous page)

B. Time-varying covariates

Covariate	Category	Wave I (2016–17)		Wave II (2017–18)		Wave III (2018–19)	
		N = 9307	100 (%)	N = 9307	100 (%)	N = 9307	100 (%)
Weekly allowance	Unknown	1539	16.5	1373	14.8	1138	12.2
	Zero	1911	20.5	1505	16.2	1166	12.5
	$1–$20	3315	35.6	2404	25.8	1632	17.5
	$21–$100	1859	20.0	2271	24.4	2411	25.9
	$100+	605	6.5	1707	18.3	2892	31.1
	Missing	78	0.8	47	0.5	68	0.7
# of physically active friends	None	484	5.2	583	6.3	729	7.8
	1 friend	962	10.3	1093	11.7	1171	12.6
	2 friends	1539	16.5	1813	19.5	1929	20.7
	3 friends	1461	15.7	1455	15.6	1531	16.5
	4 friends	767	8.2	698	7.5	731	7.9
	5 friends or more	3970	42.7	3542	38.1	3119	33.5
	Missing	124	1.3	123	1.3	97	1.0
Eating breakfast	No	4549	48.9	4812	51.7	5231	56.2
	Yes	4758	51.1	4495	48.3	4076	43.8
# of smoking friends	None	7397	79.5	6978	75.0	6904	74.2
	1 friend	998	10.7	1154	12.4	1213	13.0
	2 friends	432	4.6	548	5.9	561	6.0
	3 friends	165	1.8	234	2.5	259	2.8
	4 friends	68	0.7	78	0.8	78	0.8
	5 or more friends	147	1.6	254	2.7	236	2.5
	Missing	100	1.1	61	0.7	56	0.6
Support quit drug/alcohol	Very supportive	1791	19.2	1333	14.3	1162	12.5
	Supportive	4077	43.8	3426	36.8	3185	34.2
	Unsupportive	2324	25.0	2910	31.3	3048	32.8
	Very unsupportive	746	8.0	1273	13.7	1541	16.6
	Missing	369	4.0	365	3.9	371	4.0
# of classes skipped	0 classes	7457	80.1	6773	72.8	5754	61.8
	1 or 2 classes	1206	13.0	1583	17.0	2089	22.5
	3–5 classes	309	3.3	508	5.5	808	8.7
	6–10 classes	114	1.2	144	1.6	284	3.0
	11–20 classes	25	0.3	50	0.5	95	1.0
	More than 20 classes	26	0.3	35	0.4	74	0.8
	Missing	170	1.8	214	2.3	203	2.2
BMI category	Healthy Weight	5018	53.9	5316	57.1	5593	60.1
	Underweight	199	2.1	161	1.7	165	1.8
	Overweight	1083	11.6	1158	12.4	1237	13.3
	Obese	485	5.2	569	6.1	616	6.6
	Missing	2522	27.1	2103	22.6	1696	18.2
Gambling online	Yes	N/A		178	1.9	228	2.5
	No	N/A		8729	93.8	8684	93.3
	Missing	N/A		400	4.3	395	4.2
School connectedness	Range of [6, 24]	19.0 ± 2.9 (Mean ± SD)		18.6 ± 3.2 (Mean ± SD)		18.3 ± 3.3 (Mean ± SD)	
	Missing	157	1.7	224	2.4	233	2.5
Sedentary time (minute)	Range of [0, 2925]	425.8 ± 287.6 (Mean ± SD)		443.7 ± 297.9 (Mean ± SD)		446.3 ± 301.4 (Mean ± SD)	
	Missing	26	0.3	29	0.3	48	0.5
CESD	Range of [0, 30]	N/A		8.4 ± 5.9 (Mean ± SD)		9.1 ± 6.0 (Mean ± SD)	
	Missing	N/A		1085	11.7	943	10.1
DERS	Range of [6, 30]	N/A		14.1 ± 4.8 (Mean ± SD)		14.4 ± 4.8 (Mean ± SD)	
	Missing	N/A		481	5.2	440	4.7
GAD7	Range of [0, 21]	N/A		6.3 ± 5.6 (Mean ± SD)		6.7 ± 5.7 (Mean ± SD)	
	Missing	N/A		540	5.8	502	5.4
FLOURISH	Range of [8, 40]	N/A		32.2 ± 5.5 (Mean ± SD)		31.9 ± 5.5 (Mean ± SD)	
	Missing	N/A		261	2.8	297	3.2

(Table 1 continues on next page)

(Continued from previous page)
C. Prevalence of substance use

Substance	Category	Wave I (2016–17) N = 9307	Wave I (2016–17) 100 (%)	Wave II (2017–18) N = 9307	Wave II (2017–18) 100 (%)	Wave III (2018–19) N = 9307	Wave III (2018–19) 100 (%)
Cigarette	Never use	8320	89.4	7650	82.2	7043	75.7
	Occasional use	641	6.9	1057	11.4	1543	16.6
	Current use	283	3.0	558	6.0	672	7.2
	Missing	63	0.7	42	0.4	49	0.5
E-Cigarette	Never use	7576	81.4	6275	67.4	4636	49.8
	Occasional use	963	10.3	1263	13.6	1596	17.2
	Current use	633	6.8	1698	18.2	3008	32.3
	Missing	135	1.5	71	0.8	67	0.7
Alcohol	Never use	5637	60.6	3897	41.9	2714	29.2
	Occasional use	1902	20.4	2614	28.1	2841	30.5
	Current use	1640	17.6	2725	29.3	3678	39.5
	Missing	128	1.4	71	0.7	74	0.8
Marijuana	Never use	8192	88.0	7108	76.4	5851	62.9
	Occasional use	579	6.2	1243	13.4	1883	20.2
	Current use	406	4.4	887	9.5	1494	16.1
	Missing	130	1.4	69	0.7	79	0.8

A: Baseline descriptive at Wave I (2016–17); B: Time-varying covariates across the three waves; C: Prevalence of each substance used by type and wave. BMI: body mass index, CESD: the Center for Epidemiological Studies-Depression, DERS: the Difficulties in Emotion Regulation Scale, GAD7: the Generalized Anxiety Disorder 7-item Scale, FLOURISH: the Flourishing Scale. See Supplementary Table S1 for detailed descriptions of each variable. N/A: no measures on the COMPASS questionnaire. [a]Grades 7 & 8 are in Quebec only. [b]See Supplementary Table S1 for the definition of urban/rural classification, under "Urbanity."

Table 1: **Descriptive statistics of the three-year linked samples (N = 9307).**

samples and the prevalence of each substance used by type and wave.

The overall trend of prevalence of each substance used shows that, in general, the prevalence of "never use" had been decreasing over time, while that of "occasional use" and "current use" had been increasing for all substances across all waves. Of note, the prevalence of current use had increased significantly for e-cigarettes (4.9 times from 6.5%→32.1%) and marijuana consumption (~4 times from 4.0%→15.9%).

Dynamics of PSU patterns

Overall, four distinct PSU patterns were identified and summarised as follows: no-use of any substances (S1); occasional single-use of alcohol (S2); dual-use of e-cigarettes and alcohol (S3); and current multi-use (S4). Fig. 1 illustrates the averaged transition probability matrix across the three-wave (A), and transition matrices $W_I \rightarrow W_{II}$ (B) and $W_{II} \rightarrow W_{III}$ (C). Although the subgroup prevalence at different time occasions was similar, and the transition probability matrix revealed that an individual's use pattern membership at any time was likely to be the same as the previous time occasion, there was nevertheless a *change* between subgroups. For instance, on average, those in S2 had a 45% chance of transitioning to S3, representing the largest probability of *change* over time. In contrast, the least possible change occurred S3→S2, with the averaged transition probability being <0.01.

Fig. 2 illustrates the estimated marginal distribution of the four PSU patterns across time. It shows that the probability of S1 constantly decreased across the three-wave (0.60→0.39→0.25); the probability of S2 increased from $W_I \rightarrow W_{II}$ (0.21→0.24) and then decreased from $W_{II} \rightarrow W_{III}$ (0.24→0.22). The marginal distribution of S3 (0.14→0.25→0.33) and S4 (0.05→0.12→0.20) steadily increased over time, indicating a general tendency towards increasing use for dual and multiple substances. It is observed that the growth rate of S3 (Δ = +0.11) was 1.57 times greater than that of S4 (Δ = +0.07) from $W_I \rightarrow W_{II}$, and the growth rate for S3 (Δ = +0.08) and S4 (Δ = +0.08) was the same from $W_{II} \rightarrow W_{III}$.

By examining the incremental change (Δ) in transition probabilities from $W_{II} \rightarrow W_{III}$ vs. $W_I \rightarrow W_{II}$, we found that the probability of staying in S4 increased (Δ_{S4} = +0.08) across time. In contrast, the probability of staying in any of the lower use pattern subgroups S1 to S3 decreased over time (Δ_{S1} = −0.03, Δ_{S2} = −0.01, and Δ_{S3} = −0.02). In terms of *change*, the following transition probabilities increased across time: S1→S3 ($\Delta_{S1\rightarrow S3}$ = +0.02), S1→S4 ($\Delta_{S1\rightarrow S4}$ < +0.01), S2→S3 ($\Delta_{S2\rightarrow S3}$ = +0.01), and S3→S4 ($\Delta_{S3\rightarrow S4}$ = +0.02). On the contrary, the decreased transition probabilities included S1→S2 ($\Delta_{S1\rightarrow S2}$ < −0.01), S2→S1 ($\Delta_{S2\rightarrow S1}$ < −0.01), S2→S4 ($\Delta_{S2\rightarrow S4}$ < −0.01), S3→S1 ($\Delta_{S3\rightarrow S1}$ < −0.01), S4→S1 ($\Delta_{S4\rightarrow S1}$ = −0.05), S4→S2 ($\Delta_{S4\rightarrow S2}$ < −0.01), and S4→S3 ($\Delta_{S4\rightarrow S3}$ = −0.02). The transition probability of S3→S2 across the three-wave was unchanged

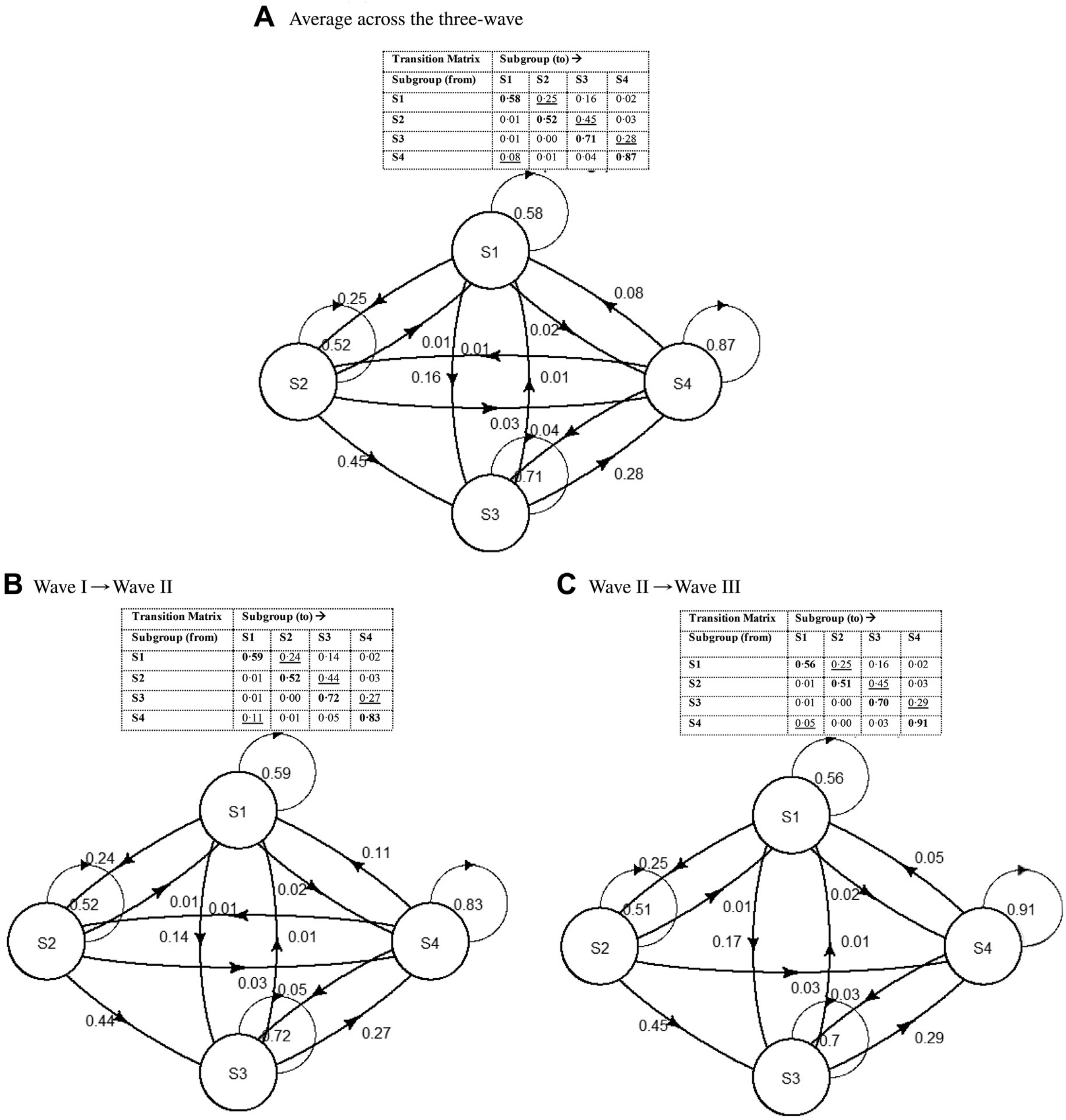


***Fig. 1:* Diagram of transition probabilities.** S1: no-use, S2: single-use of alcohol, S3: dual-use of e-cigarettes and alcohol, S4: multi-use. A: Averaged transition probabilities across the three-wave; B: Transition probabilities Wave I→Wave II; C: Transition probabilities Wave II→Wave III. Each table right on top of the transition diagram lists the corresponding transition probabilities. The diagonal in bold font indicates the largest transition probabilities in each subgroup; except for the diagonal, the second-largest transition probabilities under each subgroup were marked with an underscore.

($\Delta_{S3\rightarrow S2}$ = 0). Supplementary Table S3 summarizes these incremental changes in the initial membership probabilities over time.

Fig. 3 presents the transition patterns for each individual across the three-wave, with a table on its right summarising the prevalence of each use pattern at different time occasions, based on local decoding, i.e., the maximum posterior probability. Noted that the prevalence of S1 through S4 gradually decreased at W_{II}, being 39.0%, 24.4%, 24.9%, and 11.7%. A similar trend was observed in W_{III} data, except for S3. The longitudinal evidence of use patterns showed that although the no-use (S1) subgroup at W_I was prominent, its prevalence decreased over time ($W_I \rightarrow W_{II}$: Δ_{S1} = −21.5%; $W_{II} \rightarrow W_{III}$: Δ_{S1} = −14.0%). In contrast, the prevalence of the other three use patterns (S2 to S4) increased

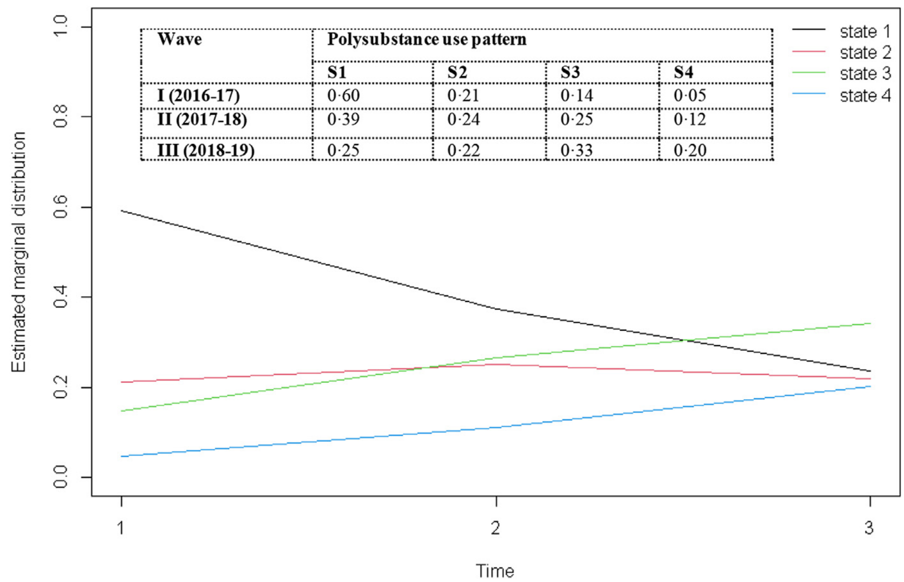


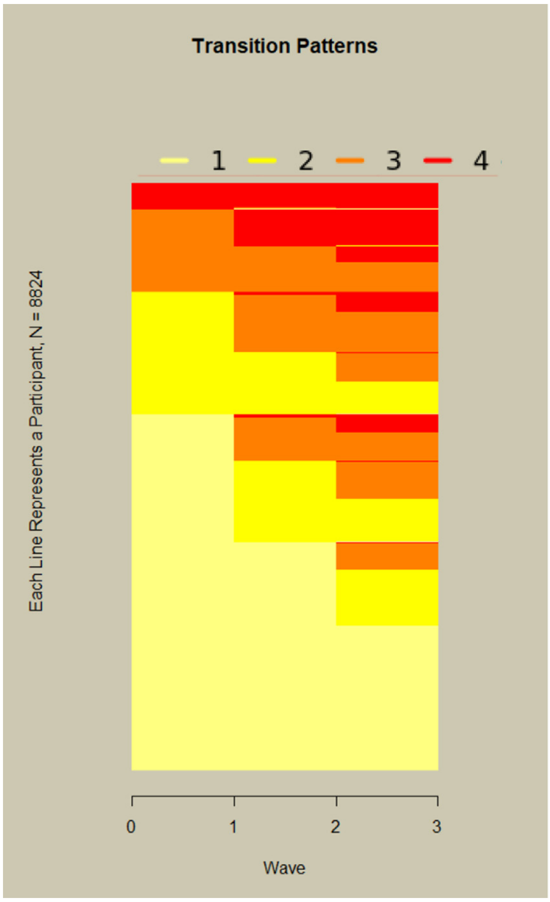
***Fig. 2:* Estimated marginal distribution of the four polysubstance use patterns (S1–S4).** Each line represents a use pattern, i.e., state 1 = S1 (no use), state 2 = S2 (single-use of alcohol), state 3 = S3 (dual-use of e-cigarettes and alcohol), and state 4 = S4 (multi-use). X-axis: three waves, Time 1 = Wave I (2016–17), Time 2 = Wave II (2017–18), Time 3 = Wave III (2018–19); Y-axis: estimated marginal distribution (values are presented in the built-in table). For example, 0.60 (S1, Wave I) means that the probability that a student belongs to the S1 (no use) subgroup at Wave I is p(S1) = 0.60.

($W_{I}\rightarrow W_{II}$: Δ_{S2} = +3.5%, Δ_{S3} = +10.9%, Δ_{S4} = +7.1%; $W_{II}\rightarrow W_{III}$: Δ_{S2} = −2.1%, Δ_{S3} = +7.6%, Δ_{S4} = +8.5%), except for S2 decreased by 2.1% from $W_{II}\rightarrow W_{III}$. By W_{III}, S3 became the prominent use pattern with the highest prevalence (32.5%). Although S4 had been the minor use pattern across the three-wave, it is alarming that the prevalence increased by 4.4 times from $W_{I}\rightarrow W_{III}$ and became very close to S2 and S1 by W_{III}.

Factors that lead to transitions

The odds ratios (ORs) for all covariates of transition between different PSU patterns are summarized in Table 2, demonstrating the average effect of each covariate on the transition probability to other use patterns, conditional on the use pattern membership at W_{I}.

Overall, being older ($OR_{S1\rightarrow S2}$ = 0.96, 95% CI [0.90, 1.03]; $OR_{S1\rightarrow S3}$ = 0.97, 95% CI [0.88, 1.07]; $OR_{S1\rightarrow S4}$ = 0.86, 95% CI [0.60, 1.12]; $OR_{S2\rightarrow S3}$ = 0.95, 95% CI [0.85, 1.05]; $OR_{S2\rightarrow S4}$ = 0.68, 95% CI [0.41, 0.95]; $OR_{S3\rightarrow S4}$ = 0.91, 95% CI [0.80, 1.03]), black (vs. white; $OR_{S1\rightarrow S2}$ = 0.90, 95% CI [0.86, 0.95], p < 0.00001; $OR_{S1\rightarrow S3}$ = 0.99, 95% CI [0.93, 1.04]; $OR_{S1\rightarrow S4}$ = 0.54, 95% CI [−0.07, 1.16]; $OR_{S2\rightarrow S3}$ = 0.97, 95% CI [0.91, 1.03]; $OR_{S2\rightarrow S4}$ = 0.94, 95% CI [0.67, 1.21]; $OR_{S3\rightarrow S4}$ = 0.82, 95% CI [0.74, 0.90], p < 0.00001), and eating breakfast ($OR_{S1\rightarrow S2}$ = 0.77, 95% CI [0.63, 0.92], p < 0.00001; $OR_{S1\rightarrow S3}$ = 0.52, 95% CI [0.31, 0.73], p < 0.00001; $OR_{S1\rightarrow S4}$ = 0.41, 95% CI [−0.45, 1.26]; $OR_{S2\rightarrow S3}$ = 0.66, 95% CI [0.45, 0.86], p < 0.001; $OR_{S2\rightarrow S4}$ = 0.88, 95% CI [0.08, 1.68]; $OR_{S3\rightarrow S4}$ = 0.59, 95% CI [0.31, 0.86], p < 0.001) were less likely to transition from low→high use direction consistently over time. Students with more weekly allowance ($OR_{S1\rightarrow S2}$ = 1.09, 95% CI [1.06, 1.12], p < 0.00001; $OR_{S1\rightarrow S3}$ = 1.11, 95% CI [1.07, 1.15], p < 0.00001; $OR_{S1\rightarrow S4}$ = 1.13, 95% CI [0.97, 1.29]; $OR_{S2\rightarrow S3}$ = 1.06, 95% CI [1.01, 1.11]; $OR_{S2\rightarrow S4}$ = 1.04, 95% CI [0.85, 1.23]; $OR_{S3\rightarrow S4}$ = 1.16, 95% CI [1.10, 1.23], p < 0.00001), more smoking friends ($OR_{S1\rightarrow S2}$ = 1.16, 95% CI [1.04, 1.27]; $OR_{S1\rightarrow S3}$ = 1.63, 95% CI [1.53, 1.73], p < 0.00001; $OR_{S1\rightarrow S4}$ = 3.07, 95% CI [2.86, 3.28]; $OR_{S2\rightarrow S3}$ = 1.40, 95% CI [1.24, 1.55], p < 0.001; $OR_{S2\rightarrow S4}$ = 2.98, 95% CI [2.72, 3.24], p < 0.00001; $OR_{S3\rightarrow S4}$ = 2.24, 95% CI [2.07, 2.40], p < 0.00001), longer sedentary time ($OR_{S1\rightarrow S2}$ = 1.02, 95% CI [1.01, 1.03]; $OR_{S1\rightarrow S3}$ = 1.06, 95% CI [1.04, 1.09], p < 0.00001; $OR_{S1\rightarrow S4}$ = 1.10, 95% CI [1.04, 1.17], p < 0.001; $OR_{S2\rightarrow S3}$ = 1.05, 95% CI [1.02, 1.07], p < 0.001; $OR_{S2\rightarrow S4}$ = 1.09, 95% CI [1.02, 1.15]; $OR_{S3\rightarrow S4}$ = 1.04, 95% CI [1.02, 1.07], p < 0.001), and attended school unsupportive ($OR_{S1\rightarrow S2}$ = 1.26, 95% CI [1.18, 1.34], p < 0.00001; $OR_{S1\rightarrow S3}$ = 1.24, 95% CI [1.12, 1.35], p < 0.001; $OR_{S1\rightarrow S4}$ = 1.92, 95% CI [1.49, 2.34], p < 0.001; $OR_{S2\rightarrow S3}$ = 1.15, 95% CI [1.03, 1.27]; $OR_{S2\rightarrow S4}$ = 1.53, 95% CI [1.08, 1.97]; $OR_{S3\rightarrow S4}$ = 1.04, 95% CI [0.90, 1.18]) were more likely to transition from low→high use direction consistently. All other covariates, including sex, urbanity, # of physically active friends, BMI category, school connectedness, and gambling online, had inconsistent effects on transition probabilities from a low→high use direction.



	Count	%
Wave I (2016-17)		
S1	5336	60·5
S2	1845	20·9
S3	1238	14·0
S4	405	4·6
Wave II (2017-18)		
S1	3445	39·0
S2	2155	24·4
S3	2193	24·9
S4	1031	11·7
Wave III (2018-19)		
S1	2204	25·0
S2	1969	22·3
S3	2868	32·5
S4	1783	20·2
Grand Total	**8824**	**100·0**

***Fig. 3:* Transition patterns for each individual across the three waves.** Reflects longitudinal trajectories of the dynamics of polysubstance use patterns over time; colours 1–4 correspond to use pattern S1–S4; each horizontal line represents a participant (N = 8824). S1: no-use, S2: single-use of alcohol, S3: dual-use of e-cigarettes and alcohol, S4: multi-use.

Likewise, the ORs on the lower-triangular matrix indicate the effects on transition probability from a high→low use direction, conditional on the reference group at W_I. Only the # of classes skipped ($OR_{S2\rightarrow S1}$ = 0.64, 95% CI [0.57, 0.71], $p < 0.00001$; $OR_{S3\rightarrow S1}$ = 0.38, 95% CI [0.33, 0.44], $p < 0.00001$; $OR_{S3\rightarrow S2}$ = 0.49, 95% CI [0.49, 0.49], $p < 0.00001$; $OR_{S4\rightarrow S1}$ = 0.00, 95% CI [−0.03, 0.04], $p < 0.00001$; $OR_{S4\rightarrow S2}$ = 0.82, 95% CI [0.17, 1.48]; $OR_{S4\rightarrow S3}$ = 0.82, 95% CI [0.65, 1.00]) was consistently associated with an increased risk of dynamic transitioning in this direction with time. All other covariates had inconsistent effects on the transition probabilities from the high→low use direction.

Discussion

This study employs ML methods to examine PSU transitions using longitudinal health survey data, revealing four distinct patterns among PSU in our large sample of youth, including no-use (S1), occasional single-use of alcohol (S2), dual-use of e-cigarettes and alcohol (S3), and current multi-use (S4), investigating the dynamics of these patterns over time and the impact factors. The evidence suggests that youth are most likely to remain in the same subgroup of use pattern or transition to a higher use group as they grow older,[4,5] which is in line with our results. We found that S4 was the most stable use pattern, with the highest probability of staying in this subgroup across time, followed by S3 and S1. S2 was the least stable pattern among these four patterns, with the lowest probability of remaining in this subgroup over time. When they transitioned, it was typically to a higher-use pattern adjacent to their current subgroup instead of a lower one, except for S4. This finding is consistent with existing literature that examines adolescent PSU using latent transition analysis. A similar trend was observed by investigating the longitudinal evidence of the transition patterns, i.e., $W_I\rightarrow W_{II}$ and $W_{II}\rightarrow W_{III}$. In particular, the chance of staying in S4 from $W_{II}\rightarrow W_{III}$ was higher than $W_I\rightarrow W_{II}$. For the other three use patterns

Characteristics/subgroup	S1	S2	S3	S4
Urbanity: "rural" (REF) vs. "small urban" vs. "medium urban" vs. "large urban" Effect of urbanity on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	0.87 (0.80–0.95)**	0.87 (0.76–0.98)	0.78 (0.36–1.20)
S2	1.59 (0.75–2.44)	REF	0.92 (0.81–1.03)	1.18 (0.71–1.65)
S3	0.59 (0.15–1.03)	1.33 (1.33–1.33)***	REF	0.87 (0.73–1.00)
S4	10.26 (9.77–10.75)***	0.57 (0.37–0.77)***	0.01 (−0.19 to 0.20)***	REF
Grade: every one-grade increase (grades 7–10 at baseline) Effect of grade/age on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	0.96 (0.90–1.03)	0.97 (0.88–1.07)	0.86 (0.60–1.12)
S2	1.24 (0.86–1.61)	REF	0.95 (0.85–1.05)	0.68 (0.41–0.95)
S3	1.21 (0.67–1.75)	0.40 (0.40–0.40)***	REF	0.91 (0.80–1.03)
S4	0.23 (−0.34 to 0.80)***	0.37 (−0.07 to 0.82)**	0.57 (0.19–0.95)**	REF
Race/Ethnicity: "White" (REF) vs. "Black" vs. "Asian" vs. "Aboriginal (First Nations, Métis, Inuit)" vs. "Latin American/Hispanic" vs. "Other" Effect of race/ethnicity on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	0.90 (0.86–0.95)***	0.99 (0.93–1.04)	0.54 (−0.07 to 1.16)
S2	1.18 (0.81–1.56)	REF	0.97 (0.91–1.03)	0.94 (0.67–1.21)
S3	1.24 (0.84–1.63)	1.60 (1.60–1.60)***	REF	0.82 (0.74–0.90)***
S4	1.13 (0.63–1.63)	1.26 (0.86–1.66)	0.04 (−0.02 to 0.10)***	REF
Weekly allowance: "I do not know how much money I get each week" (REF) vs. "Zero" vs. "$1–$20" vs. "$21–$100" vs. "$100+" Effect of weekly allowance on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	1.09 (1.06–1.12)***	1.11 (1.07–1.15)***	1.13 (0.97–1.29)
S2	0.39 (−0.22 to 1.00)**	REF	1.06 (1.01–1.11)	1.04 (0.85–1.23)
S3	0.72 (0.34–1.10)	0.98 (0.98–0.98)***	REF	1.16 (1.10–1.23)***
S4	1.17 (0.78–1.56)	2.29 (2.04–2.55)***	0.49 (0.00–0.97)**	REF
# of physically active friends: "None" (REF) vs. "1 friend" vs. "2 friends" vs. "3 friends" vs. "4 friends" vs. "5 friends or more" Effect of # of physically active friends on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	1.23 (1.19–1.28)***	1.24 (1.18–1.31)***	1.28 (1.05–1.52)
S2	0.60 (0.03–1.18)	REF	1.23 (1.16–1.29)***	0.93 (0.67–1.18)
S3	0.90 (0.35–1.44)	1.67 (1.67–1.67)***	REF	1.08 (1.00–1.17)
S4	0.28 (−0.16–0.72)***	1.06 (0.52–1.60)	16.06 (15.96–16.16)***	REF
Eating breakfast: "No" (REF) vs. "Yes" Effect of eating breakfast on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18 For example, $OR_{S1 \rightarrow S3}$ = 0.52 means that the odds for the event (i.e., transitioning from the no-use subgroup S1 at Wave I to the dual-use of e-cigarettes and alcohol subgroup S3 at Wave II) in those who reported eating breakfast were 0.52 times the odds in the comparison group (i.e., those who reported not eating breakfast).				
S1	REF	0.77 (0.63–0.92)**	**0.52 (0.31–0.73)*****	0.41 (−0.45 to 1.26)
S2	1.12 (0.99–1.26)	REF	0.66 (0.45–0.86)**	0.88 (0.08–1.68)
S3	1.30 (1.18–1.42)***	1.46 (1.46–1.46)***	REF	0.59 (0.31–0.86)**
S4	6.46 (6.36–6.55)***	0.42 (0.37–0.47)***	92874.09 (92874.02–92874.16)***	REF
# of smoking friends: "None" (REF) vs. "1 friend" vs. "2 friends" vs. "3 friends" vs. "4 friends" vs. "5 or more friends" Effect of # of smoking friends on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				

(Table 2 continues on next page)

Characteristics/subgroup	S1	S2	S3	S4
(Continued from previous page)				
For example, $OR_{S1 \rightarrow S4}$ = 3.07 means that the odds for the event (i.e., transitioning from the no-use subgroup S1 at Wave I to the multi-use subgroup S4 at Wave II) in those who reported having one smoking friend were 3.07 times the odds in the comparison group (i.e., those who reported no friends who smoke). The same OR applies to all categorical comparisons, i.e., "None" vs. "1 friend" vs. "2 friends" vs. "3 friends" vs. "4 friends" vs. "5 or more friends."				
S1	REF	1.16 (1.04–1.27)	1.63 (1.53–1.73)***	**3.07 (2.86–3.28)***
S2	0.46 (0.36–0.55)***	REF	1.40 (1.24–1.55)**	2.98 (2.72–3.24)***
S3	0.44 (0.39–0.49)***	0.62 (0.62–0.62)***	REF	2.24 (2.07–2.40)***
S4	0.00 (0.00–0.01)***	1.12 (0.69–1.54)	0.00 (−0.01 to 0.01)***	REF
Support quit drug/alcohol: "Very supportive" (REF) vs. "Supportive" vs. "Unsupportive" vs. "Very unsupportive" Effect of support quit drug/alcohol on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	1.26 (1.18–1.34)***	1.24 (1.12–1.35)**	1.92 (1.49–2.34)**
S2	4.08 (3.73–4.42)***	REF	1.15 (1.03–1.27)	1.53 (1.08–1.97)
S3	0.11 (−0.09 to 0.32)***	0.53 (0.53–0.53)***	REF	1.04 (0.90–1.18)
S4	0.19 (−0.50 to 0.87)***	1.06 (0.26–1.87)	0.33 (0.14–0.52)***	REF
Sex: "Female" (REF) vs. "Male" Effect of sex on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18 For example, $OR_{S3 \rightarrow S1}$ = 1.54 means that the odds for the event (i.e., transitioning from the dual-use of e-cigarettes and alcohol subgroup S3 at Wave I to the no-use subgroup S1 at Wave II) in males were 1.54 times the odds in the comparison group (i.e., females).				
S1	REF	0.70 (0.56–0.84)***	1.38 (1.18–1.58)**	2.17 (1.35–2.99)
S2	0.94 (0.90–0.97)**	REF	1.12 (0.91–1.32)	0.22 (−0.14 to 0.58)***
S3	**1.54 (1.50–1.59)***	0.99 (0.99–0.99)***	REF	1.88 (1.62–2.13)***
S4	66.11 (65.99–66.23)***	2.02 (1.93–2.11)***	2.32 (2.26–2.37)***	REF
# of classes skipped: "0 classes" (REF) vs. "1 or 2 classes" vs. "3 to 5 classes" vs. "6 to 10 classes" vs. "11 to 20 classes" vs. "More than 20 classes" Effect of # of classes skipped on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	1.27 (1.17–1.38)***	1.71 (1.59–1.82)**	2.43 (2.16–2.71)
S2	0.64 (0.57–0.71)***	REF	1.35 (1.21–1.49)**	2.41 (2.11–2.72)***
S3	0.38 (0.33–0.44)***	0.49 (0.49–0.49)***	REF	1.51 (1.36–1.65)***
S4	0.00 (−0.03 to 0.04)***	0.82 (0.17–1.48)	0.82 (0.65–1.00)	REF
BMI category: "Healthy Weight" (REF) vs. "Underweight" vs. "Overweight" vs. "Obese" vs. "Not Stated" Effect of BMI category on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	0.89 (0.84–0.93)***	0.88 (0.82–0.94)**	0.84 (0.60–1.07)
S2	0.53 (0.16–0.90)**	REF	0.95 (0.88–1.02)	0.88 (0.61–1.14)
S3	2.06 (1.38–1.75)	1.01 (1.01–1.01)***	REF	1.01 (0.93–1.09)
S4	1.13 (0.63–1.63)	2.35 (1.67–3.03)	16.84 (16.28–17.40)***	REF
School connectedness: every one-unit increase in score (ranging from 6 to 24) Effect of school connectedness on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	1.02 (1.00–1.05)	0.94 (0.91–0.97)**	0.92 (0.81–1.03)
S2	0.75 (0.50–1.00)	REF	1.01 (0.97–1.04)	0.90 (0.78–1.02)
S3	1.10 (0.86–1.35)	0.87 (0.87–0.87)***	REF	0.95 (0.91–1.00)
S4	1.11 (0.85–1.38)	1.11 (0.88–1.34)	0.92 (0.68–1.17)	REF
Sedentary time: every 1-h increase Effect of sedentary time on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18) Conditional on … (vertical: use pattern at baseline, W_{I}, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	1.02 (1.01–1.03)	1.06 (1.04–1.09)***	1.10 (1.04–1.17)**
S2	0.90 (0.73–1.11)	REF	1.05 (1.02–1.07)**	1.09 (1.02–1.15)
S3	0.90 (0.69–1.18)	0.43 (0.43–0.44)***	REF	1.04 (1.02–1.07)**
S4	0.68 (0.50–0.93)	0.93 (0.79–1.09)	1.36 (1.17–1.58)**	REF

(Table 2 continues on next page)

Characteristics/subgroup	S1	S2	S3	S4
(Continued from previous page)				
Gambling online: "Yes" (REF) vs. "No" Effect of gambling online on the probability of transitioning to … (horizontal: use pattern at W_{II}, 2017–18)→ Conditional on … (vertical: use pattern at baseline, W_I, 2016–17) Relative to being in the same use pattern at W_{II}, 2017–18				
S1	REF	1.40 (1.00–1.79)	0.88 (0.37–1.39)	0.16 (−0.01 to 0.33)***
S2	1.15 (1.08–1.22)**	REF	1.32 (0.80–1.85)	3.30 (3.24–3.36)***
S3	0.66 (0.60–0.71)***	0.66 (0.66–0.66)***	REF	1.41 (0.90–1.93)
S4	290.62 (290.56–290.67)***	0.37 (0.31–0.43)***	0.26 (0.18–0.35)***	REF

Note: ***$p < .00001$; **$p < .001$. All covariates entered simultaneously as predictors of use pattern membership at baseline Wave I (2016–17) and Wave I (2016–17)→Wave II (2017–18) transition. Bold values indicate interpretations. The OR on the upper-triangular matrix indicates the effects on transition probability from a low→high use direction, conditional on the comparison group (REF) at Wave I. The OR on the lower-triangular matrix indicates the effects on transition probability from a high→low use direction, conditional on the comparison group (REF) at Wave I. S1: no use; S2: single-use of alcohol; S3: dual-use of e-cigarettes and alcohol; S4: multi-use. 95% confidence intervals are in brackets.

Table 2: **Odds ratios and 95% confidence intervals for all predictors of transition between PSU patterns (N = 8824).**

(S1 through S3), the stability decreased over time, implying that students starting at any of these use patterns had an increased chance of transitioning to other use patterns over time.

In terms of *change* from W_I→W_{III}, the most likely transition occurred S2→S3, followed by S3→S4, and S1→S2. In contrast, students in S3 were least likely to transition to S2, followed by S4→S2, S2→S1, and S3→S1. Similar to the longitudinal observation of the transition probabilities for *stability*, we examined the incremental change in transition probabilities across the three-wave. In general, the chances of transitioning from a low→high use direction increased over time. On the contrary, the decreased transition probabilities indicate slimmer chances of moving from a high→low use direction with time, except for S1→S2 and S2→S4.

Not only do use patterns change with time, but so does the evidence outside this study about use patterns. For example, legalisation of recreational cannabis use in Canada in 2018 may explain increased self-report at W_{III} (2018–19) without the fear of legal consequences, plus the availability and easier access to cannabis products on the legal market. With the emerging trend of e-cigarette use among youth, including e-cigarettes as new evidence while examining use patterns would be more meaningful than ever. Unfortunately, no previous studies have shown that exploring the dynamics of youth PSU patterns also uses e-cigarettes as an indicator of substance use. While this is a novelty of the current study, it also makes the direct comparison between our findings and other evidence challenging.

In terms of the factors that lead to the dynamics in PSU patterns, Choi et al.[4] (2018) reported that males were more likely to transition from legal to more illicit substance use than females, while female polysubstance users were more likely to transition to a less use pattern than males. Our finding of the sex difference on the dynamic transition of use patterns partly agrees with Choi et al., i.e., male students were more likely to transition to a higher use group over time, except for transitioning from S1→S2 and S2→S4. However, we found that males were more likely to transition to a lower use pattern as well than their female peers, except for transitioning from S2→S1 and S3→S2. Except for the sex difference, there is inadequate literature about other variables leading to the dynamics of PSU pattern membership. Thanks to rich longitudinal evidence available in COMPASS data, we examined multifaceted covariates to determine if they were significant in predicting the dynamic transitions of use patterns over time. Their effects are more complex than those on the initial membership of use patterns across the four use patterns across the three waves. These covariates range from demographic information to health behaviours, from individual- to population-level.

Truancy, being measured by the # of classes skipped, significantly affects transition probabilities from any low→high and high→low use directions over time. Generally, other risk factors that lead to bi-directional dynamic transitions of use patterns include having more weekly allowance (except for transitioning from S4→S1 and S4→S2), more smoking friends (except for transitioning from S4→S2), longer sedentary time (except for transitioning from S4→S3), attending school unsupportive (except for transitioning from S2→S1, S4→S2). Similarly, the odds for transitioning from a low→high use direction in students who reported not gambling online were >1 times the odds in those who were gambling online, except for S1→S3 and S1→S4, and the odds for transitioning from a high→low use direction in those not gambling online were <1 times the odds in online gamblers, except for S2→S1 and S4→S1. Students who belonged to these subgroups were more likely to transition to a higher-use group and were less likely to transition to a lower-use group with time than those in corresponding comparison groups.

On the other hand, in general, protective factors that are associated with bi-directional dynamics include being black (vs. white, except for transitioning from S4→S3), eating breakfast (except for transitioning from

S4→S2), and unhealthy weight (vs. healthy weight, except for transitioning from S2→S1 and S3→S4). In other words, students in these subgroups were less likely to transition to a higher-use group and were more likely to transition to a lower-use group over time than those in corresponding comparison groups.

In addition to the sex difference, age, urbanity, school connectedness, and the # of physically active friends have inconsistent effects on the transition probabilities, particularly from a high→low use direction. For example, when students age, the odds for transitioning from a low→high use direction in older students were <1 times the odds in younger students, and the odds for transitioning from the opposite use direction (high→low) in older students were also found to be <1 times the odds in younger students, except for S2→S1 and S3→S1. Students living in a rural area, having more physically active friends were more likely to transition to a higher-use group (except for transitioning from S2→S4), and were found with mixed effects on the transition probabilities from a high→low use direction.

Our study results have several public health implications. First, the dynamics of PSU patterns in adolescence can inform school board programming design on how to deal with relevant health threats at this developmental stage and throughout the process. An early detection-prevention approach could be initiated with a more effective strategy for the low or intermediate-use pattern groups, particularly the least stable pattern S2. Second, harm reduction programs targeting the multi-use pattern group S4 may help these high-risk students learn coping strategies, improve health behaviours, make a positive change, and prevent more costly substance abuse treatment later in their lives. Third, our results indicate that dual-use of e-cigarettes and alcohol has become the most common cohort as adolescents age. Although policymakers' priorities for addressing lower use pattern groups (S1 and S2) and high-risk students (S4) are recommended, services targeting this cohort (S3) and breaking the S3→S4 transition is critical. Furthermore, public health practitioners should pay more attention to modifiable factors identified in this study while designing and implementing any quit smoking/alcohol/drugs programs, which should not be a stand-alone practice. Instead, school policies should integrate these initiatives with other approaches like fostering physical activity, healthy eating, anti-depression, etc.

The strengths of this study lie in the COMPASS dataset and the methodologies we applied, including the LASSO regression, a superior approach for feature selection, MI techniques to impute missing values, and the LMM modelling method to evaluate transitioning of latent processes corresponding to health behaviours without standard measurement protocols. LMMs were initially developed in multiple fields with applications in sociology, psychology, and medicine,[32] e.g., examining the tendency of substance use.[33] This modelling technique can be applied to any content area with similar longitudinal data to address more social science research questions that include complicated transitions across time, such as mental health and behaviour change, and can better inform the management and treatment of addiction and other health issues.

However, this study has certain limitations. The first one is the limited number of waves in our dataset, which hinders our ability to provide multi-level granularity for transition modelling on PSU patterns. The second limitation is the lack of external validation data to evaluate our model's performance. Furthermore, all responses to the Cq are self-reported, including substance use measures. Thus, the precision of the measures used in reporting substance use is subject to self-reported bias and may lead to imprecise estimates. Moreover, the factors included in examining the dynamic transitions in use patterns are limited within the scope of the Cq. Although the family history of substance use can be a key factor in youth's exposure and acceptance of substance use, such data elements are unavailable in the Cq. Similarly, no question asks about the sibling information of the participants in the Cq. If there is more than one child per family in the same school, each one was treated as an individual participant in these analyses. Lastly, we need to apply caution when interpreting and generalising results since many participating schools in the COMPASS study are purposefully sampled and may not be genuinely representative. For example, most participants in this study came from Ontario (67.5%), which limits the external validity of the study results. Omitting other Canadian provinces may limit generalizability, as youth PSU may differ in these regions. For instance, youth living in Northern provinces have the highest per-capita rate of hospitalization involving marijuana and alcohol consumption in the country.[34]

Of note, younger age groups (Grades 7–8) are only available in Quebec, with their own norms distinct from the rest of the provinces in and beyond the linked samples. Students in Grades 7–8 and 9–10 differ significantly concerning multiple factors, such as independence, more financial ability to purchase substances, expansion of social networks, transition from middle/elementary school to high school, more experience and easier access to substances, and longer duration of substance use as one age. Any of these factors may contribute to increased PSU as adolescents age. Future work is warranted to explore age differences in more detail and analyze other characteristic differences in the dynamic transition of use patterns among youth by conducting a stratified analysis using the LMM method. Likewise, mental illness is known to be associated with substance use. There is one dedicated section in the Cq asking about students' mental health from W_{II} onwards.

However, after running the LASSO regression, none of the mental health factors was selected for further model building. We plan to add these data elements and those reflecting school environments, school health policies and practices to the LMM model to observe the dynamics of PSU over a longer period.

Conclusions

The current PSU trend among adolescents has become a growing challenge facing many countries with severe consequences both for the individual and our society. As the first study to ascertain the dynamics of use patterns and the factors that lead to the transition in youth PSU using COMPASS data, we demonstrate the application of LMMs in settings with complex and high-dimensional population-level longitudinal health survey data. This study provides evidence for the opportunities and possibilities for using these methods to improve the prevention and management of substance abuse issues. An aspiration behind this study is to provide a means to accelerate the joint research that can provide insights into designing and implementing programs and interventions for those directly affected by the detrimental effects of youth PSU. Findings from this study can be beneficial to practitioners in the field, such as school program managers or policymakers, in their capacity to develop interventions to prevent or remedy PSU where these distinct profiles can be considered in the design and deployment.

Contributors

Yang Yang: conceptualisation, data curation, formal analysis, methodology, project administration, visualisation, writing – original draft, and writing – review & editing. Zahid A. Butt: conceptualisation, funding acquisition, methodology, project administration, resources, supervision, validation, and writing – review & editing. Scott T. Leatherdale: conceptualisation, funding acquisition, resources, supervision, validation, and writing – review & editing. Plinio P. Morita: supervision, and writing – review & editing. Alexander Wong: methodology, writing – review & editing. Laura Rosella: resources, and writing – review & editing. Helen H. Chen: conceptualisation, funding acquisition, methodology, project administration, resources, supervision, validation, and writing – review & editing.

Data sharing statement

COMPASS study data is available upon request through the completion and approval of an online form: https://uwaterloo.ca/compass-system/information-researchers/data-usage-application. The datasets used during the current study are available from the corresponding author upon reasonable request.

Declaration of competing interests

We declare no competing interests.

Acknowledgments

Y.Y. was supported by the Applied Health Sciences (AHS) scholarship and Microsoft AI for Good grant. The COMPASS study has been supported by a bridge grant from the CIHR Institute of Nutrition, Metabolism and Diabetes (INMD) through the "Obesity – Interventions to Prevent or Treat" priority funding awards (OOP-110788; awarded to S.L.), an operating grant from the CIHR Institute of Population and Public Health (IPPH) (MOP-114875; awarded to S.L.), a CIHR project grant (PJT-148562; awarded to S.L.), a CIHR bridge grant (PJT-149092; awarded to S.L.), a CIHR project grant (PJT-159693), and by a research funding arrangement with Health Canada (#1617-HQ-000012; contract awarded to S.L.), and a CIHR-Canadian Centre on Substance Abuse (CCSA) team grant (OF7 B1-PCPEGT 410-10-9633; awarded to S.L.). The COMPASS-Quebec project additionally benefits from funding from the Ministère de la Santé et des Services sociaux of the province of Québec, and the Direction régionale de santé publique du CIUSSS de la Capitale-Nationale.

We acknowledge the assistance received from Fulvia Pennoni, Professor of Statistics from the Department of Statistics and Quantitative Methods at the University of Milano-Bicocca, and Francesco Bartolucci, Professor of Statistics from the Department of Economics at the University of Perugia in Italy, in the use of the LMest package for LMM modelling.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lana.2022.100389.

References

1 Silveira ML, Green VR, Iannaccone R, Kimmel HL, Conway KP. Patterns and correlates of polysubstance use among US youth aged 15–17 years: wave 1 of the population assessment of tobacco and health (PATH) study. *Addiction*. 2019;114(5):907–916.
2 Bray BC, Smith RA, Piper ME, Roberts LJ, Baker TB. Transitions in smokers' social networks after quit attempts: a latent transition analysis. *Nicotine Tob Res*. 2016;18(12):2243–2251.
3 Halladay J, Woock R, El-Khechen H, et al. Patterns of substance use among adolescents: a systematic review. *Drug Alcohol Depend*. 2020;216:108222.
4 Choi HJ, Lu Y, Schulte M, Temple JR. Adolescent substance use: latent class and transition analysis. *Addict Behav*. 2018;77:160–165.
5 Connor JP, Gullo MJ, White A, Kelly AB. Polysubstance use: diagnostic challenges, patterns of use and health. *Curr Opin Psychiatry*. 2014;27(4):269–275.
6 Zuckermann AM, Williams G, Battista K, de Groh M, Jiang Y, Leatherdale ST. Trends of poly-substance use among Canadian youth. *Addict Behav Rep*. 2019;10:100189.
7 Zuckermann AM, Williams GC, Battista K, Jiang Y, de Groh M, Leatherdale ST. Prevalence and correlates of youth poly-substance use in the COMPASS study. *Addict Behav*. 2020;107:106400.
8 Zuckermann AM, Battista KV, Bélanger RE, et al. Trends in youth cannabis use across cannabis legalization: data from the COMPASS prospective cohort study. *Prev Med Rep*. 2021;22:101351.
9 Zuckermann AM, Gohari MR, Romano I, Leatherdale ST. Changes in cannabis use modes among Canadian youth across recreational cannabis legalization: data from the COMPASS prospective cohort study. *Addict Behav*. 2021;122:107025.
10 Haines-Saah RJ, Fischer B. Youth cannabis use and Legalization in Canada–reconsidering the fears, myths and facts three years in. *J Can Acad Child Adolesc Psychiatry*. 2021;30(3):191.
11 Tomczyk S, Isensee B, Hanewinkel R. Latent classes of polysubstance use among adolescents—a systematic review. *Drug Alcohol Depend*. 2016;160:12–29.
12 Merrin GJ, Leadbeater B. Do classes of polysubstance use in adolescence differentiate growth in substances used in the transition to young adulthood? *Subst Use Misuse*. 2018;53(13):2112–2124.
13 Canadian Tobacco, Alcohol and Drugs Survey (CTADS): summary of results for 2017 [internet]. Available from: https://www.canada.ca/en/health-canada/services/canadian-alcohol-drugs-survey/2017-summary.html; 2019. Accessed March 8, 2021.
14 Strunin L, Díaz-Martínez A, Díaz-Martínez LR, et al. Age of onset, current use of alcohol, tobacco or marijuana and current polysubstance use among male and female Mexican students. *Alcohol Alcohol*. 2017;52(5):564–571.

15 Patton GC, Coffey C, Carlin JB, Degenhardt L, Lynskey M, Hall W. Cannabis use and mental health in young people: cohort study. *BMJ*. 2002;325(7374):1195–1198.
16 Mathers M, Toumbourou JW, Catalano RF, Williams J, Patton GC. Consequences of youth tobacco use: a review of prospective behavioural studies. *Addiction*. 2006;101(7):948–958.
17 Lesjak V, Stanojević-Jerković O. Physical activity, sedentary behavior and substance use among adolescents in Slovenian urban area/Telesna Aktivnost, Oblike Sedečega Vedenja In Uživanje Psihoaktivnih Snovi Med Mladostniki V Slovenskem Urbanem Okolju. *Slovenian J Public Health*. 2015;54(3):168–174.
18 Isralowitz RE, Trostler N. Substance use: toward an understanding of its relation to nutrition-related attitudes and behavior among Israeli high school youth. *J Adolesc Health*. 1996;19(3):184–189.
19 White J, Walton D, Walker N. Exploring comorbid use of marijuana, tobacco, and alcohol among 14 to 15-year-olds: findings from a national survey on adolescent substance use. *BMC Public Health*. 2015;15(1):1–9.
20 Su J, Supple AJ, Kuo SI. The role of individual and contextual factors in differentiating substance use profiles among adolescents. *Subst Use Misuse*. 2018;53(5):734–743.
21 Pettigrew S, Jongenelis M, Lawrence D, Rikkers W. Common and differential factors associated with abstinence and poly drug use among Australian adolescents. *Int J Drug Policy*. 2017;50:41–47.
22 Tomczyk S, Hanewinkel R, Isensee B. Multiple substance use patterns in adolescents—a multilevel latent class analysis. *Drug Alcohol Depend*. 2015;155:208–214.
23 Cranford JA, McCabe SE, Boyd CJ. Adolescents' nonmedical use and excessive medical use of prescription medications and the identification of substance use subgroups. *Addict Behav*. 2013;38(11):2768–2771.
24 Hale D, Viner R. Trends in the prevalence of multiple substance use in adolescents in England, 1998–2009. *J Public Health*. 2013;35(3):367–374.
25 Leatherdale ST, Brown KS, Carson V, et al. The COMPASS study: a longitudinal hierarchical research platform for evaluating natural experiments related to changes in school-level programs, policies and built environment resources. *BMC Public Health*. 2014;14(1):1–7.
26 Reel B, Bredin C, Leatherdale ST. COMPASS year 5 and 6 school recruitment and retention compass technical report series. In: *Compass technical report series*. 2018.
27 Thompson-Haile A, Leatherdale ST. Student-level data collection procedures. In: *Compass technical report series*. 2013.
28 Kaplan D. An overview of Markov chain methods for the study of stage-sequential developmental processes. *Dev Psychol*. 2008;44(2):457.
29 Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B Methodol*. 1996;58(1):267–288.
30 R Core Team. *R: a language and environment for statistical computing*. 2013.
31 Bartolucci F, Farcomeni A, Pandolfi S, Pennoni F. LMest: an R package for latent Markov models for categorical longitudinal data. *arXiv*. 2015. preprint arXiv:1501.04448.
32 Bartolucci F, Farcomeni A, Pennoni F. *Latent Markov models for longitudinal data*. CRC Press; 2012.
33 Bartolucci F. Likelihood inference for a class of Latent Markov Models under linear hypotheses on the transition probabilities. *J R Stat Soc Ser B Stat Methodol*. 2006;68(2):155–178.
34 N.W.T. youth have highest rate of hospitalization for substance abuse in country: study. Laura Busch, CBC News, September 19, 2019. Available from: https://www.cbc.ca/news/canada/north/n-w-t-youth-have-highest-rate-of-hospitalization-for-substance-abuse-in-country-study-1.5289230. Accessed July 7, 2022.