

Initiation and persistence of health risk behaviors through adolescence: longitudinal findings from urban South Africa

Alysse Kowalski

Emory University Laney Graduate School <https://orcid.org/0000-0001-7200-9694>

O. Yaw Addo

Emory University School of Public Health

Michael R Kramer

Emory University Department of Epidemiology

Reynaldo Martorell

Emory University School of Public Health

Shane A Norris

University of the Witwatersrand Faculty of Health Sciences

Rachel N Waford

Emory University School of Public Health

Linda M Richter

University of the Witwatersrand Faculty of Health Sciences

Aryeh Stein (✉ aryeh.stein@emory.edu)

<https://orcid.org/0000-0003-1138-6458>

Research article

Keywords: Adolescence, risk behavior, low- and middle-income country, smoking, alcohol, cannabis, illicit drug use, sexual activity, adolescent pregnancy

Posted Date: March 16th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-17353/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published on January 11th, 2021. See the published version at <https://doi.org/10.1186/s12887-020-02486-y>.

Abstract

Background: Little is known about longitudinal patterns of adolescent health risk behavior initiation and persistence in low- and middle-income countries.

Methods: Birth to Twenty Plus is a longitudinal birth cohort in Soweto-Johannesburg, South Africa. We used reports from Black African participants on smoking, alcohol, cannabis, illicit drugs, and sexual activity and on adolescent pregnancy collected over 7 study visits between ages 11 and 18 y. We fit Kaplan-Meier curves to estimate behavior initiation or adolescent pregnancy, examined current behavior at age 18 y by age of initiation, and performed a clustering analysis to identify patterns of initiation and their sociodemographic predictors.

Results: By age 13 y, cumulative incidence of smoking and alcohol initiation were each >21%, while the cumulative incidence of other behaviors and adolescent pregnancy were <5%. By age 18 y, smoking, alcohol, and sexual activity initiation estimates were each >65%, cannabis (age 15 y) and illicit drug initiation were each >16%; adolescent pregnancy was 31%. At both ages, rates of initiation were higher among males. At age 18 y, current risk behavior activity was lower than lifetime activity and generally unrelated to age of initiation. We identified three clusters reflecting low, moderate, and high-risk patterns of risk behavior initiation. One-third of males and 17% of females were assigned to the high-risk cluster. Sociodemographic factors were not associated with cluster membership.

Conclusions: Among urban dwelling Black South Africans, risk behavior experimentation across adolescence is common and clusters into distinct initiation patterns unrelated to the sociodemographic factors assessed. Understanding patterns of risk behavior initiation has implications for the timing of primary and secondary public health interventions and supports integrated prevention efforts that consider multiple behaviors simultaneously.

Background

During adolescence, defined by the WHO as ages 10–19 y, the development of the reward and pleasure centers in the brain contributes to risk-taking and sensation-seeking, rendering a degree of experimentation a normative attribute of adolescence (1, 2). Six of the ten leading risk factors of morbidity and mortality among young people ages 15–19 y, and three of the ten among young people age 10 to 14 y, are behavioral, including smoking, alcohol use, drug use, and unsafe sex (3). When established in adolescence, behavioral risk factors have consequences that extend into adulthood. Early tobacco use increases the risk of regular tobacco and cannabis use, use of hard drugs and drug problems, alcohol problems, and early pregnancy (4–7). Earlier age of first alcohol use predicts alcohol abuse and lifetime dependence and other substance use (5, 8–12). Earlier sexual debut increases the number of sexual partners and risk of pregnancy and STIs including HIV (13).

There is a longstanding interest in risk behavior co-occurrence, but far less is known about patterns of risk behavior initiation. This is important because different risk behaviors are initiated at different times; some may act as gateways to other behaviors, and the acceptability of risk behaviors changes over adolescence. To better understand patterns of risk behavior initiation across adolescence requires longitudinal data to consider multiple risk behaviors simultaneously to identify subgroups of individuals with similar patterns of risk behavior engagement (14–16). Previous work has generally examined either multiple behaviors at a single time point or a single behavior at multiple time points. Studies examining multiple behaviors at a single time point cannot account for changes in risk behavior co-occurrence over time (14, 15). Studies examining a single behavior over multiple time points often use latent class growth analysis to provide a detailed understanding of that behavior but do not account for co-occurrence (17–19).

This work has several implications for public health policy makers and practitioners. Improved understanding of patterns of risk behavior initiation would allow for better timing of primary prevention efforts intended to prevent risk behavior initiation and secondary prevention efforts to mitigate risk behaviors once begun.

Despite being home to 90% of the world's adolescents, there is a paucity of longitudinal data on adolescent health from low and middle-income countries (20). To improve our understanding of health risk behaviors among adolescents in urban South Africa, we describe longitudinal patterns of smoking, alcohol, cannabis, illicit drug, and sexual activity initiation and of adolescent

pregnancy; examine current behavior at age 18 y by stage of initiation; and use clustering analysis to characterize patterns of risk behavior initiation.

Methods

Birth to Twenty Plus (Bt20+) cohort

Birth to Twenty Plus (Bt20+) is an observational birth cohort in Soweto-Johannesburg, South Africa. The study enrolled singleton children born between April and June 1990 who resided in the municipal area for a minimum of 6 months after birth (N = 3273). Almost 70% of cohort members were still traceable when they were age 17 y, with the majority of attrition occurring during the preschool years (21).

Ethical approval

Ethical clearance for this study was provided by the University of the Witwatersrand Human Research Ethics Committee for Research on Human Subjects (M181186) and the Emory University Institutional Review Board (00062989).

Data collection

We used data from 7 waves of data collected in adolescence (we refer to these as the age 11, 13, 14, 15, 16, 17, and 18 y study visits) and pregnancy data through age 18 y from ongoing pregnancy and live birth surveillance. Study visits were completed at the Developmental Pathways for Health Research Unit at Chris Hani Baragwanath Hospital in Soweto. Data were collected by interview or using self-administered questionnaires. Self-administered questionnaires were completed on paper at the Year 11, 13, and 14 study visits, and using an audio computer-aided self-administered interview (CASI) system at ages 15, 16, 17, and 18 y.

Risk behaviors

The risk behaviors of interest include smoking, alcohol use, cannabis use, illicit drug use, and sexual activity. Questions were adapted from the US and South African Youth Risk Behavior Surveys, with additional items developed specifically for the study (22, 23). Questions on smoking and sexual activity were asked at all adolescent study visits; alcohol use at ages 11, 13, and 18 y; cannabis use from age 11 to 15 y; and illicit drug use at ages 11, 13, 14, 15, 17 and 18 y. For each risk behavior, data were captured on three aspects: 1) risk behavior initiation; 2) age of initiation; and 3) use or activity in the past month as of the age 18 y study visit.

Risk behavior initiation

We defined risk behavior initiation as an affirmative response to a Yes/No question about ever engaging in a behavior. We defined *illicit drug use initiation* as an affirmative response to at least one question about ever use of five drugs for which repeated measures were available (inhalants/glue, ecstasy, mandrax (Quaaludes), cocaine, or LSD) at ages 11, 13, 14, and 15 y or an affirmative response to any lifetime drug use at the age 17 and 18 y study visits.

Age of initiation

We defined *age of initiation* using the age of first engagement reported by the Bt20+ participant. If this was unavailable, we used the individual's age at the time of the study visit or, if this could not be calculated from the date of the visit and the date of birth, we assigned the age corresponding to the year of study visit (Supplemental Table 1). Age of initiation was not asked for cannabis and illicit drug use. For these behaviors we assigned the respondent's exact age or the age corresponding to the study year the first time these were reported. For all behaviors we set implausible ages of initiation (ages > 2 years above the study visit OR ages < 5 y) to missing. We categorized individuals as initiating a behavior in childhood (<11 y), early (11–13 y), mid (14–16 y), or late adolescence (17–18 y), never at age 18 y, and status unknown at age 18 y (2).

Current activity at age 18 y

We defined current activity at 18 y as an affirmative response to a Yes/No question about use or activity in the past 30 days.

Adolescent pregnancy

We defined adolescent pregnancy as an affirmative response to the pregnancy history question, first asked at age 15 y, or having report of pregnancy through age 18 y in the surveillance system. We defined age of adolescent pregnancy using the age captured in the surveillance system, the respondent's exact age at the study visit, or the age corresponding to the study year, in that order.

Sociodemographic characteristics

Maternal age at birth, years of schooling, and marital status were collected at enrollment into the study. We used tertiles of the number of household assets owned as a measure of socioeconomic position in early life (using data from age 0–2 y) and in childhood (using data from age 7 y supplemented with data from age 5 y).

At the age 0–2, 5, and 7 y study visits mothers were asked about stress and violence events experienced in the past six months. We characterized childhood exposure to stress as the number of study waves at which the mother reported more than the sample median number of events.

Data management and preparation

Analytical sample

We excluded cohort members from non-Black African population groups (22% of the cohort) who comprise less than 10% of the population in Soweto-Johannesburg. Of the 2568 Black African participants enrolled in the cohort, 1822 attended at least one study visit during adolescence and contributed information for at least one risk behavior of interest. We excluded 82 individuals who reported sexual activity prior to age 12 y from the sexual activity and pregnancy analyses as this was before the legal age of consent. To maximize sample sizes, we retained participants with data on any of the measures of interest in the analytical dataset; therefore, the sample sizes vary by measure (Supplemental Figure 1, Supplemental Table 2).

Risk behavior descriptive analyses

We fit Kaplan-Meier curves for each behavior to estimate the probability of reaching age 19 y without initiating that behavior. Individuals who did not report an event were censored using their age at the last study visit they attended. We examined current behavior activity at age 18 y and used chi-square tests to examine associations between stage of risk behavior initiation and current activity at 18 y.

Cluster analysis and risk behavior profiles

We conducted a hierarchical agglomerative cluster analysis among 1,126 individuals for whom status was known for all risk behaviors. We used Gower's method to calculate the dissimilarity matrix between individuals based on stage of initiation and applied Ward's method to evaluate the similarity between clusters and determine which clusters to combine at each iteration, and calculated a series of fit indices using the NbClust R package (24, 25). For females, a three-cluster solution was indicated by a majority of the fit statistics. Although a two-cluster solution was indicated by the same criterion among males, the addition of a third cluster meaningfully differentiated an additional subgroup of adolescents by subdividing the first cluster into two, without crossover from the other cluster (Supplemental Tables 3). We examined associations of sociodemographic characteristics with cluster membership.

Sensitivity analysis

We compared demographic characteristics of individuals included in the analytical sample to those who were excluded to assess potential bias due to attrition prior to adolescence, nonresponse to the risk behavior questions, or reported sexual activity before the age of consent. We compared individuals included in the cluster analysis to individuals with incomplete information to assess selection bias in the cluster analysis.

All analyses were sex-specific and conducted using R version 3.5.3 (26). We considered two-sided p-values <0.05 statistically significant.

Results

Sample characteristics

Study participant's mothers were in their mid-twenties and had 9.58 (2.74) years of schooling on average at enrollment, and 66% were single (Table 1). Among Black African participants, those included in the study included in the study were born to mothers with an additional year of schooling and who were more likely to be single than those who were excluded. Most excluded individuals were lost to follow-up early in the study and therefore have no information from later study waves.

Table 1
Demographic characteristics of Black African Birth to Twenty Plus participants by inclusion status^a

	Included (n = 1822)	Excluded (n = 746)	p-value ^b
Sex - Males	880 (48%)	369 (49%)	0.62
- Females	942 (52%)	377 (51%)	
Maternal age at birth	25.81 (6.26)	25.95 (5.87)	0.59
Maternal years of schooling	9.58 (2.74)	8.53 (3.59)	< 0.01
Marital status - Single	1206 (66%)	425 (57%)	< 0.01
- Partnered	608 (34%)	320 (43%)	
Asset tertile in early life - 1	624 (34%)	287 (38%)	< 0.01
- 2	362 (20%)	87 (12%)	
- 3	512 (28%)	76 (10%)	
- Missing	324 (18%)	296 (40%)	
Asset tertile at age 7 y - 1	664 (36%)	78 (10%)	< 0.01
- 2	398 (22%)	33 (4%)	
- 3	466 (26%)	45 (6%)	
- Missing	294 (16%)	590 (79%)	
Child stress in early life - Below median	498 (27%)	210 (28%)	0.02
- Above median	401 (22%)	129 (17%)	
- Missing	923 (51%)	407 (55%)	
Child stress at age 5 y - Below median	750 (41%)	79 (11%)	< 0.01
- Above median	441 (24%)	28 (4%)	
- Missing	631 (35%)	639 (86%)	
Child stress at age 7 y - Below median	921 (51%)	87 (12%)	< 0.01
- Above median	542 (30%)	35 (5%)	
- Missing	359 (20%)	624 (84%)	
^a Presented as N (%) or mean ± SD.			
^b Chi-square or t-test p-values.			

Kaplan-Meier curves for the probability of “surviving” adolescence without initiating a risk behavior are summarized in Figure 1. By the end of adolescence (age 18 y), estimates of smoking, alcohol, and sexual activity initiation each exceeded 75% and illicit drug use exceeded 30% among males. Patterns were similar among females, but rates of initiation were slightly lower. By age 13 y, estimates of smoking and alcohol initiation were 41.6% and 34.5% among males and 21.1% and 23.1% among females. Among both sexes, cannabis use was predominantly initiated in mid-adolescence and drug use initiated in mid- and late adolescence.

Current risk behavior engagement at age 18 y and stage of initiation

Current substance use and sexual activity reported at age 18 y was lower than lifetime use for all behaviors (Table 2). Individuals who first used illicit drugs in late adolescence were twice as likely to report current drug use at age 18 y as compared to individuals who first used illicit drugs earlier in adolescence (64% late vs 33% early among males and 37% late vs 11% early among females) (Table 2). Males who started smoking in late adolescence were less likely to be current smokers at age 18 y than males who started smoking in early adolescence (29% late vs 52% early).

Table 2
Risk behavior activity at age 18 y by stage of initiation^a

Measure	N initiate behavior	N current behavior	Childhood	Early	Mid	Late	p-value ^b
Males							
Smoking	719	278	18 (44%)	130 (52%)	115 (53%)	15 (29%)	0.01
Alcohol use	537	300	2 (33%)	132 (59%)	109 (72%)	57 (56%)	0.02
Illicit drug use	273	89	NA	4 (33%)	22 (22%)	62 (64%)	< 0.01
Sexual activity	577	222	NA	40 (52%)	135 (54%)	47 (41%)	0.05
Females							
Smoking	649	140	8 (50%)	37 (25%)	71 (27%)	24 (23%)	0.15
Alcohol use	521	210	2 (18%)	62 (38%)	85 (49%)	61 (42%)	0.05
Illicit drug use	133	30	NA	1 (11%)	7 (16%)	21 (37%)	0.04
Sexual activity	629	293	NA	3 (50%)	136 (61%)	154 (52%)	0.10
^a Displayed as N (%) of individuals who initiated a behavior in a given stage that reported current behavior activity at age 18 y.							
^b Chi-square p-values testing differences in stage of initiation among current users.							

Cluster analysis and risk behavior profiles

We identified a three-cluster solution in which the clusters represent different patterns of risk behavior initiation in adolescence and reflect groups of adolescents with low, moderate, and high-risk patterns based on initiation in a given stage of adolescence within a cluster compared to overall initiation (Figure 2). The high-risk cluster (33% of males and 17% of females) includes individuals who reported risk behavior initiation at rates higher than the group mean. Individuals initiating illicit drug or cannabis use in early or mid-adolescence were almost exclusively classified in the high-risk cluster. Individuals in the moderate risk cluster (33% of males and 60% of females) initiated smoking, alcohol, and sexual activity at rates higher than the group mean in adolescence and did not use cannabis. The remaining individuals were in the low-risk cluster (33% of males and 23% of females) and initiated risk behaviors at rates below the group mean. When these individuals did initiate a behavior, it tended to be in late adolescence. For example, none of the females in the low-risk cluster reported smoking in early or mid-adolescence and 15% reported initiating smoking in late adolescence in comparison to the mean group initiation rates of 20% in early, 37% in mid, and 14% in late adolescence. Females in the moderate and high-risk clusters had similar rates of sexual initiation by age 18 y, though females in the high-risk cluster were more likely to initiate sex in mid-adolescence. Rates of adolescent pregnancy were higher

among females in the high-risk cluster compared to females in the low and moderate risk clusters (47% high vs 35% moderate and 19% low). None of the sociodemographic factors examined were associated with cluster membership (Table 3).

Table 3

Bivariate associations of selected sociodemographic characteristics with patterns of risk behavior initiation^a

	Males				Females			
	Low risk (n = 169)	Moderate risk (n = 159)	High risk (n = 172)	p-value ^b	Low risk (n = 131)	Moderate risk (n = 342)	High risk (n = 98)	p-value ^b
Maternal years of schooling	9.58 (2.6)	9.87 (2.6)	9.61 (2.4)	0.53	9.7 (2.5)	9.64 (3.0)	9.8 (2.4)	0.89
Maternal age at birth	26.08 (6.4)	25.43 (6.3)	26.17 (6.7)	0.54	26.13 (6.2)	25.85 (6.4)	26.15 (6.5)	0.87
Marital status - Single/Separated/Divorced	110 (65%)	105 (66%)	100 (58%)	0.26	83 (63%)	234 (69%)	62 (63%)	0.37
- Partnered	59 (35%)	54 (34%)	72 (42%)		48 (37%)	105 (31%)	36 (37%)	
Asset tertile in early life - 1	58 (41%)	53 (42%)	64 (46%)	0.76	49 (43%)	115 (41%)	25 (32%)	0.06
- 2	28 (20%)	30 (24%)	27 (19%)		37 (33%)	78 (28%)	19 (24%)	
- 3	55 (39%)	42 (34%)	48 (35%)		27 (24%)	90 (32%)	35 (44%)	
Early life assets imputed - 0	145 (86%)	145 (91%)	146 (85%)	0.18	107 (82%)	290 (85%)	83 (85%)	0.70
- 1	24 (14%)	14 (9%)	26 (15%)		24 (18%)	52 (15%)	15 (15%)	
Asset tertile at age 7 y - 1	67 (47%)	55 (42%)	56 (39%)	0.58	50 (44%)	115 (38%)	34 (42%)	0.06
- 2	34 (24%)	36 (28%)	46 (32%)		33 (29%)	84 (28%)	12 (15%)	
- 3	41 (29%)	39 (30%)	41 (29%)		30 (27%)	102 (34%)	34 (42%)	
Age 7 asset source - assets from Y7	140 (99%)	129 (99%)	140 (98%)	0.66	108 (96%)	284 (94%)	75 (94%)	0.84
- assets from Y5	2 (1%)	1 (1%)	3 (2%)		5 (4%)	17 (6%)	5 (6%)	
Child stress - Above median stressful events never	73 (46%)	67 (46%)	66 (42%)	0.43	43 (36%)	140 (43%)	41 (46%)	0.54
- Above median stressful events 1X	50 (31%)	57 (39%)	56 (36%)		48 (40%)	120 (37%)	33 (37%)	
- Above median stressful events 2 or 3X	36 (23%)	23 (16%)	35 (22%)		29 (24%)	67 (20%)	15 (17%)	
Number of stress measures - Attended 1 study visit	26 (16%)	36 (24%)	23 (15%)	0.19	19 (16%)	60 (18%)	19 (21%)	0.88

^a Presented as N (%) or mean ± SD.^b Chi-square p-values testing differences across patterns of risk behavior initiation.

	Males		Females		
- Attended 2 study visits	73 (46%)	57 (39%)	75 (48%)	59 (49%)	158 (48%) 43 (48%)
- Attended 3 study visits	60 (38%)	54 (37%)	59 (38%)	42 (35%)	109 (33%) 27 (30%)
^a Presented as N (%) or mean ± SD.					
^b Chi-square p-values testing differences across patterns of risk behavior initiation.					

Sensitivity analysis

Individuals were required to have a known status at age 18 y for the five behaviors of interest to be included in the cluster analysis. Individuals included in the cluster analysis did not differ from their excluded peers on sociodemographic characteristics including sex, household asset ownership in early life and childhood, and childhood exposure to stress and violence. Percent initiation was comparable between the included and excluded groups in childhood and early adolescence but excluded individuals more likely to have an unknown status at age 18 y as individuals were lost to follow-up over the course of adolescence (Supplemental Table 4).

Discussion

In this cohort of urban dwelling Black African adolescents, risk behavior experimentation was common and clustered into three distinct profiles that reflect low, moderate, and high-risk patterns of risk behavior initiation across adolescence. Interestingly, household sociodemographic factors did not predict profiles of risk.

Initiation of smoking, alcohol, cannabis, illicit drugs (among males), and sexual activity, and rates of pregnancy by age 18 y were substantially higher in this cohort as compared to those 18 years of age in a 2008 nationally representative cross-sectional survey of South African students (15). The prevalence of smoking initiation was 86% among males and 72% among females in Bt20+ compared to 37% among males and 18% among females in the South African Youth Risk Behaviour Survey (YRBS). In contrast, estimates of alcohol and cannabis initiation at age 18 y were lower in Bt20+ than estimates from 12th graders surveyed in the 2007 United States YRBS (when Bt20 participants were age 17 y), though lifetime smoking and sexual activity were higher (27). Cross-sectional studies are prone to recall bias (28). Additionally, risk patterning may differ by race and urbanicity. While the 2008 South African YRBS age 18 y estimates were not disaggregated by race, at all ages the prevalence of smoking and alcohol use were lower among Black Africans than other South African racial groups (White, Coloured, and Indian), while drug use and sexual activity were higher. Estimates of smoking and alcohol use were 10% and 15% higher in the mostly urban Gauteng province, where Bt20+ is located, compared to the national average.

It is unsurprising that among those who ever experimented with substance use, current use at age 18 y was much lower. In Bt20+ more than 60% of individuals who initiated smoking were no longer smoking at age 18 y whereas a greater proportion of individuals who used alcohol were still using. In the South African YRBS 30% of individuals who initiated smoking were no longer smoking at age 18 y, while 25% to 35% of individuals who initiated alcohol use were still drinking at age 18 y (15). These smaller differences may be attributable to recall bias and differences between an urban and a nationally representative sample.

We identified three distinct subgroups reflecting low, moderate, and high-risk patterns of risk behavior initiation. As clustering analyses are data driven, it is challenging to draw comparisons with other studies. The results of a cluster analysis of university students in the UK identified three clusters based on smoking and alcohol use as well as stress and lifestyle factors (14). Like the moderate and high-risk clusters in this study, one of their clusters was characterized by smoking and binge drinking, though that study did not consider illicit drug use.

In an analysis of a representative sample from the Netherlands, risk behaviors were shown to cluster differently with age. Specifically, smoking, alcohol, and drug use clustered together among adolescents age 12 to 15 y (questions about unsafe sex were not asked of this age group), while at ages 16 to 18 y unsafe sex and alcohol use clustered together, and smoking, drug use, and other delinquent behaviors clustered together (16). Individuals in our high-risk cluster initiated smoking and alcohol use at

above average rates in early adolescence and illicit drug use before age 17 y and initiated sexual activity at above average rates in early and mid-adolescence. The Dutch study identified a “healthy” cluster characterized by favorable diet and physical activity behaviors; we did not examine these behaviors in this analysis.

A cluster analysis in the 2008 South African YRBS identified low, intermediate, and high-risk clusters. Individuals in the YRBS high-risk cluster had substance use, sexual behavior, and traffic safety domain scores at least twice the national average (15). In our high-risk cluster, illicit drug use was initiated at greater than twice the national rate and sexual activity initiation was above average.

Finally, the temporal sequencing of risk behavior initiation in our high-risk cluster is consistent with other findings that have shown smoking and alcohol initiation in early adolescence are associated with subsequent cannabis and illicit drug use (6, 10). Bt20+ females in the high-risk cluster were more likely to become pregnant during adolescence, a life-altering event that limits girls’ educational and socioeconomic prospects. Associations of socioeconomic status with cluster assignment in other studies have been mixed, with higher socioeconomic status associated with both low risk cluster membership and engagement in increased number of risk behaviors (15, 29, 30). Interestingly, none of the sociodemographic characteristics considered were associated with cluster assignment. Further research is needed to identify predictors of cluster membership.

Strengths

We used longitudinal data to describe smoking, alcohol, cannabis, illicit drug use, and sexual activity initiation and adolescent pregnancy over the course of adolescence in an urban, middle-income country context underrepresented in the literature. Our study had high response rates and limited attrition during follow up. Due to the longitudinal design, we were able to describe patterns of adolescence risk behaviors initiation prospectively, thus limiting recall bias.

Limitations

Some limitations should also be considered. The analyses use self-reported data, which may introduce bias, though anonymity was assured during data collection. Furthermore, we used paper-based self-administered questionnaires until age 14 y, after which an audio-CASI was used; both approaches have demonstrated acceptable validity and reliability for sensitive subjects (31–36). Repeated measures of current use were not available consistently, though by using serial measures of lifetime use we were able to examine the age of initiation for five health risk behaviors. Though the data used in this analysis were collected from 2001 to 2009, cross-sectional surveys of youth risk behavior in South Africa from 2002 to 2011 showed little to no change in the prevalence of adolescent risk behaviors (15, 23, 37).

Public health implications

High levels of risk behavior initiation support the use of public health interventions to prevent initiation and long-term persistence. Distinct patterns in initiation inform the type of prevention efforts warranted, when they should be implemented, and which behaviors should be targeted together. By age 13 y, smoking and alcohol use prevalence were already >20% in girls and 35% in boys, therefore primary prevention efforts should be targeted to younger children. By the end of early adolescence, secondary prevention efforts to mitigate risk behavior engagement should be incorporated. As observed in the moderate and high-risk initiation patterns, smoking and alcohol use are often initiated in the same stage of adolescence (early or mid) while cannabis and illicit drug use are more likely to be initiated in mid-adolescence.

Conclusions

These data provide a valuable reference for smoking, alcohol, cannabis, drug use and sexual activity in a well-characterized cohort of Black African adolescents in Soweto-Johannesburg, South Africa to which contemporary studies can be compared. The present study clearly demonstrates high levels of risk behavior experimentation over the course of adolescence that should be addressed by public health interventions to prevent risk behavior initiation and adoption.

Abbreviations

Bt20+: Birth to Twenty Plus study

CASI: computer-aided self-administered interview

WHO: World Health Organization

Declarations

Ethics approval

Ethics approval for this study was provided by the Emory University Institutional Review Board (00062989) and University of the Witwatersrand Human Research Ethics Committee (M181186). Consent to participate in research was obtained at each study visit. Written consent by the Bt20+ participant or written consent by the caregiver and verbal assent by the Bt20+ participant were obtained as age appropriate.

Consent for publication

NA

Availability of data and materials

The datasets used in this analysis are available from Dr. Shane Norris (shane.norris@wits.ac.za) upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

We would like to acknowledge the NIH Fogarty International Center (R25TW009337), Bill & Melinda Gates Foundation (OPP1164115), Wellcome Trust (077210/Z/05/Z), the DSI-NRF Centre of Excellence in Human Development hosted by the University of the Witwatersrand, and the South African Medical Research Council for supporting this work.

Authors' contributions

AJK and ADS conceived and designed the work. AJK conducted the analysis and drafted the manuscript. OYA, MRK, RM, SAN, RNW, and LMR critically reviewed and provided substantive comments on the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would also like to thank the study participants and their families for their time and participation in the Birth to Twenty Plus cohort.

References

1. Casey BJ, Getz S, Galvan A. The adolescent brain. *Developmental review*. 2008;28(1):62–77.
2. World Health Organization. Recognizing adolescence online2014 [Defining terms]. Available from: <http://apps.who.int/adolescent/second-decade/section2/page1/recognizing-adolescence.html>.
3. Mokdad AH, Forouzanfar MH, Daoud F, Mokdad AA, El Bcheraoui C, Moradi-Lakeh M, et al. Global burden of diseases, injuries, and risk factors for young people's health during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*. 2016;387(10036):2383–401.

4. DuRant RH, Smith JA, Kreiter SR, Krowchuk DP. The Relationship Between Early Age of Onset of Initial Substance Use and Engaging in Multiple Health Risk Behaviors Among Young Adolescents. *JAMA Pediatrics*. 1999;153(3):286–91.
5. Ellickson PL, Hays RD, Bell RM. Stepping through the drug use sequence: longitudinal scalogram analysis of initiation and regular use. *Journal of abnormal psychology*. 1992;101(3):441–51.
6. Ellickson PL, Tucker JS, Klein DJ. High-risk behaviors associated with early smoking: results from a 5-year follow-up. *Journal of Adolescent Health*. 2001;28(6):465–73.
7. 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):948–58.
8. Chou SP, Pickering RP. Early onset of drinking as a risk factor for lifetime alcohol-related problems. *British journal of addiction*. 1992;87(8):1199–204.
9. Gonzalez GM. Early onset of drinking as a predictor of alcohol consumption and alcohol-related problems in college. *Journal of drug education*. 1989;19(3):225–30.
10. Gruber E, DiClemente RJ, Anderson MM, Lodico M. Early drinking onset and its association with alcohol use and problem behavior in late adolescence. *Prev Med*. 1996;25(3):293–300.
11. Kandel D, Yamaguchi K. From beer to crack: developmental patterns of drug involvement. *American journal of public health*. 1993;83(6):851–5.
12. Yu J, Williford WR. The age of alcohol onset and alcohol, cigarette, and marijuana use patterns: an analysis of drug use progression of young adults in New York State. *The International journal of the addictions*. 1992;27(11):1313–23.
13. Baily M-C, Baggaley RF, Wang L, Masse B, White RG, Hayes RJ, et al. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *The Lancet infectious diseases*. 2009;9(2):118–29.
14. Dodd LJ, Al-Nakeeb Y, Nevill A, Forshaw MJ. Lifestyle risk factors of students: A cluster analytical approach. *Preventive Medicine*. 2010;51(1):73–7.
15. Reddy S, James S, Sewpaul R, Koopman F, Funani N, Sifunda S, et al. Umthente Uhlaba Usamila—The South African Youth Risk Behaviour Survey 2008. Cape Town: South African Medical Research Council; 2010.
16. van Nieuwenhuijzen M, Junger M, Velderman MK, Wiefferink KH, Paulussen TWGM, Hox J, et al. Clustering of health-compromising behavior and delinquency in adolescents and adults in the Dutch population. *Preventive Medicine*. 2009;48(6):572–8.
17. Colder CR, Campbell RT, Ruel E, Richardson JL, Flay BR. A finite mixture model of growth trajectories of adolescent alcohol use: Predictors and consequences. *Journal of consulting and clinical psychology*. 2002;70(4):976.
18. Colder CR, Mehta P, Balanda K, Campbell RT, Mayhew K, Stanton WR, et al. Identifying trajectories of adolescent smoking: an application of latent growth mixture modeling. *Health Psychology*. 2001;20(2):127.
19. Connell CM, Gilreath TD, Hansen NB. A Multiprocess Latent Class Analysis of the Co-Occurrence of Substance Use and Sexual Risk Behavior Among Adolescents. *Journal of Studies on Alcohol and Drugs*. 2009;70(6):943–51.
20. UN Department of Social and Economic Affairs PD. World population prospects: highlights of the 2008 revision. Working Paper No ESA/P/WP.210. [Working Paper No ESA/P/WP.210]. In press 2009.
21. Norris SA, Richter LM, Fleetwood SA. Panel studies in developing countries: case analysis of sample attrition over the past 16 years within the birth to twenty cohort in Johannesburg, South Africa. *Journal of International Development: The Journal of the Development Studies Association*. 2007;19(8):1143–50.

22. Centers for Disease Control and Prevention. Youth Risk Behavior Surveillance System (YRBSS) 2018 [Available from: <https://www.cdc.gov/healthyyouth/data/yrbs/index.htm>].
23. Reddy S, Panday S, Swart D, Jinabhai C, Amosun S, James S, et al. Umthente uhlaba usamila – The South African youth risk behaviour survey 2002. Cape Town: South African Medical Research Council; 2003.
24. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*. 2014;61(6):1–36.
25. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. cluster: Cluster Analysis Basics and Extensions. R package version 2.0.8; 2019.
26. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing 2019.
27. Centers for Disease Control and Prevention. 1991–2017 High School Youth Risk Behavior Survey Data.
28. Brener ND, Billy JO, Grady WR. Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: evidence from the scientific literature. *Journal of adolescent health*. 2003;33(6):436–57.
29. Zweig JM, Lindberg LD, McGinley KA. Adolescent Health Risk Profiles: The Co-Occurrence of Health Risks among Females and Males. *Journal of Youth and Adolescence*. 2001;30(6):707–28.
30. de Winter AF, Visser L, Verhulst FC, Vollebergh WA, Reijneveld SA. Longitudinal patterns and predictors of multiple health risk behaviors among adolescents: The TRAILS study. *Prev Med*. 2016;84:76–82.
31. Beauclair R, Meng F, Deprez N, Temmerman M, Welte A, Hens N, et al. Evaluating audio computer assisted self-interviews in urban South African communities: evidence for good suitability and reduced social desirability bias of a cross-sectional survey on sexual behaviour. *BMC medical research methodology*. 2013;13:11.
32. Dolezal C, Marhefka SL, Santamaria EK, Leu CS, Brackis-Cott E, Mellins CA. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Archives of sexual behavior*. 2012;41(2):401–10.
33. Metzger DS, Koblin B, Turner C, Navaline H, Valenti F, Holte S, et al. Randomized controlled trial of audio computer-assisted self-interviewing: utility and acceptability in longitudinal studies. HIVNET Vaccine Preparedness Study Protocol Team. *American journal of epidemiology*. 2000;152(2):99–106.
34. Del Boca FK, Darkes J. The validity of self-reports of alcohol consumption: state of the science and challenges for research. *Addiction*. 2003;98 Suppl 2:1–12.
35. Smith GT, McCarthy DM, Goldman MS. Self-reported drinking and alcohol-related problems among early adolescents: dimensionality and validity over 24 months. *Journal of studies on alcohol*. 1995;56(4):383–94.
36. Strecher VJ, Becker MH, Clark NM, Prasada-Rao P. Using patients' descriptions of alcohol consumption, diet, medication compliance, and cigarette smoking: the validity of self-reports in research and practice. *Journal of general internal medicine*. 1989;4(2):160–6.
37. Reddy S, James S, Sewpaul R, Sifunda S, Ellahebokus A, Kambaran NS, et al. Umthente uhlaba usamila: The 3rd South African national youth risk behaviour survey 2011. 20

Figures

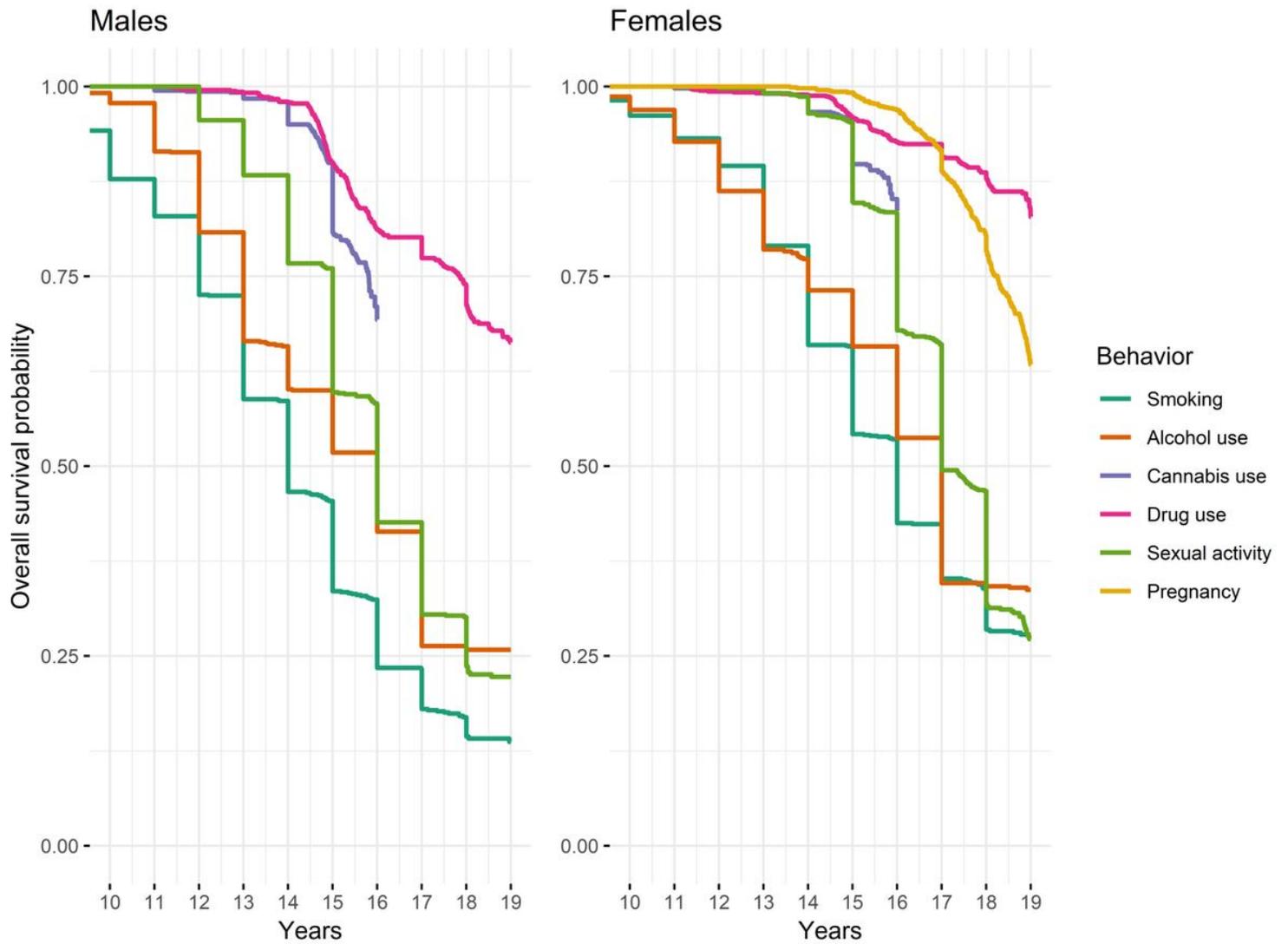


Figure 1

Cumulative survival probability of smoking, alcohol use, marijuana use, illicit drug use, and sexual activity initiation and pregnancy through age 18 years. As age of first smoke, alcohol use, and sexual activity were primarily determined from self-reported integer age, these curves follow a stepwise decline in contrast to the more gradual declines in the marijuana use, illicit drug use, and pregnancy curves, which were primarily determined from the participant's exact calendar age at the study visit.

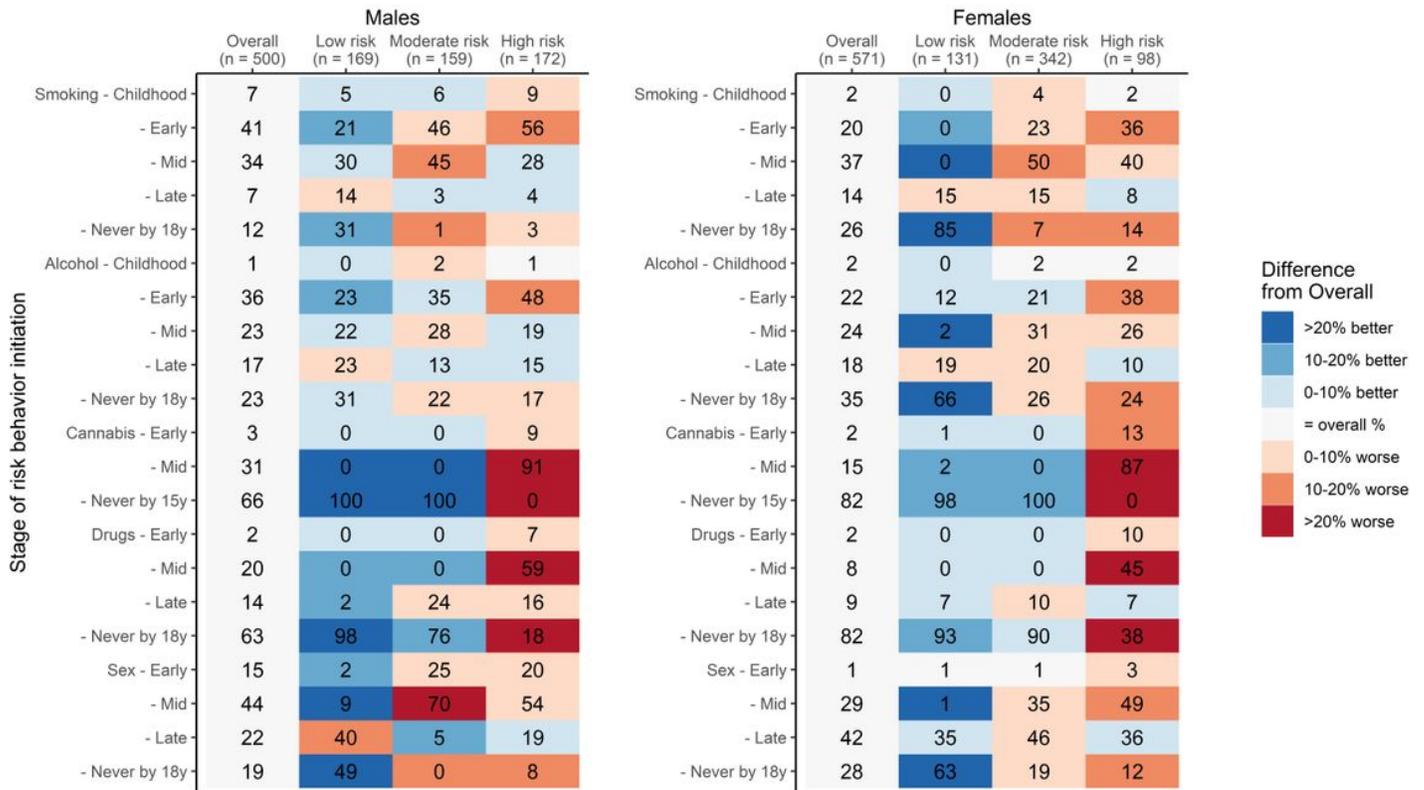


Figure 2

Patterns of percent risk behavior initiation by stage of childhood and adolescence. A cell color reflects the degree to which percent initiation in a given cluster differs from the overall study population – blue cells reflect below average initiation while red cells reflect above average percent initiation, with deeper shades reflecting greater absolute differences. “Never” initiation was reverse color-coded such that cluster percentages higher than the overall percentage reflect “better” health.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.pdf](#)