

Further investigations of gateway effects using the PATH study

Peter N Lee (✉ PeterLee@pnlee.co.uk)

John S Fry

RoeLee Statistics Ltd

Research article

Keywords: Cigarettes, confounding, over-adjustment, e-cigarettes, gateway effects, modelling, propensity score

Posted Date: March 26th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-18592/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 at F1000Research on November 9th, 2020. See the published version at <https://doi.org/10.12688/f1000research.24289.2>.

Abstract

Background: Considerable interest exists in whether e-cigarette use (“vaping”) by youths increases the risk of initiating cigarette smoking. Based on Waves 1 and 2 of the Population Assessment of Tobacco and Health study we reported that adjustment for propensity for vaping using Wave 1 variables explained about 80% of the unadjusted relationship. This analysis may be over-adjusted had vaping at Wave 1 affected some variables recorded then. Here we present analyses using Waves 1 to 3 to avoid this possibility.

Methods: Our main analysis M1 concerned those who had never smoked by Wave 2 and never vaped by Wave 1. Wave 2 vaping was linked to smoking initiation by Wave 3, adjusting for Wave 1 predictors. Sensitivity analyses excluded other tobacco product users at Wave 1, included other tobacco product use as an additional predictor, or were based on propensity for ever smoking or ever any tobacco use, rather than ever vaping. Other analyses adjusted for propensity as derived originally, or ignored Wave 1 data. Other analyses used grouped age (only available originally) or exact age (available now) as a confounder variable, attempted residual confounding adjustment by modifying values of predictor variables using data later recorded, or considered interactions with age.

Results: In M1, propensity adjustment removed about 50% of the excess odds ratio (i.e. OR–1), the unadjusted OR, 5.60 (95% CI 4.52–6.93) becoming 3.37 (2.65–4.28), 3.11 (2.47–3.92) or 3.27 (2.57–4.16) depending whether adjustment was for propensity as a continuous variable, as quintiles, or for the 16 variables making up the propensity score. Many factors studied hardly affected the results, including using grouped or exact age, consideration of other tobacco products, including interactions, or using predictors of smoking or tobacco use rather than vaping. The clearest conclusion was that analyses avoiding over-adjustment only explained about 50% of the excess OR whereas analyses subject to over-adjustment explained about 80%.

Conclusions: Although much of the unadjusted gateway effect results from uncontrolled confounding, our current analysis provides stronger evidence of a causal effect of vaping than did our earlier analysis. However, some doubts remain about the completeness of confounder adjustment.

Background

In youths, use of e-cigarettes (“vaping”) has increased considerably in recent years in many countries (e.g. [1–3]). It is generally recognized that vaping significantly reduces exposure to harmful constituents compared to smoking [4], so one might expect risks from vaping to be much lower [5]. However, there are concerns about the rise in vaping. The concern of interest here is the possibility that vaping may encourage some individuals to start smoking who would otherwise not have done so, often referred to as the “gateway” effect. The concern that vaping may act as a gateway into smoking was originally brought sharply into focus by a 2017 meta-analysis [6] based on nine US cohort studies in young people linking previous vaping to subsequent initiation of smoking. This paper reported that among baseline never-smokers, ever vaping at baseline strongly predicted initiation in the next 6 to 18 months, with an odds ratio (OR) of 3.62 (95% confidence interval (CI) 2.42–5.41) after adjustment for various predictors of initiation. Similarly baseline past 30-day vaping also predicted subsequent 30-day cigarette use (OR 4.25, 95% CI 2.52–7.37).

We have previously published two papers relating to the gateway effect. Our first paper [7] considered various general issues. It made a number of relevant points:

Although studies reported that vaping significantly predicts smoking initiation following adjustment for various other predictors, the sets of predictors considered were generally quite incomplete.

No study considered residual confounding arising from inaccurate measurement of predictors.

More precise adjustment may have substantially reduced the association.

Any true gateway effect would likely have affected smoking prevalence only modestly.

Smoking prevalence in US and UK youths in 2014–2016 declined somewhat faster than predicted by the preceding trend, whereas a substantial gateway effect would suggest the opposite.

Even were some gateway effect to exist, introducing e-cigarettes would still be likely to reduce deaths from smoking-related diseases.

Our second paper [8] described results of our own analyses, based on data from Waves 1 and 2 of the Population Assessment of Tobacco and Health (PATH) study, a nationally representative longitudinal cohort study in the United States of tobacco use and how it affects the health of people. Wave 1 was conducted from 12 September 2013 to 15 December 2014, with Wave 2 the first annual follow-up. For each Wave, data are available separately for Youths (aged 12–17 years) and Adults (aged 18+ years), the Youth data including some information from the parents. Publicly available data files include extensive information on use of various types of tobacco

products and on a range of variables linked to initiation of tobacco. Note that where youths become 18 between successive waves of the survey, their data will be available in the Adult data rather than the Youth data. Also additional youths who were under 12 at the time of Wave 1 are added into the Youth data when they reach the age of 12 at a subsequent wave.

The main analyses we described considered Wave 1 never cigarette smoking youths who, at Wave 2, had information available on smoking initiation. Having constructed a propensity score for ever e-cigarette use from Wave 1 variables, we found that adjustment markedly reduced the unadjusted OR of 5.70 (95% CI 4.33–7.50) to 2.48 (1.85–3.31), 2.47 (1.79–3.42) or 1.85 (1.35–2.53), depending on whether adjustment was made for propensity as quintiles, for propensity as a continuous variable, or for the individual variables making up the propensity score. Various sensitivity analyses confirmed that adjustment removed most of the gateway effect.

Although we found that confounding was a major factor, explaining most of the observed gateway effect, we were particularly concerned about the possibility of over-adjustment, if taking up e-cigarettes had affected the values of some of the Wave 1 predictor variables considered. At the time, we noted that the possibility of over-adjustment could be avoided using data from Waves 1, 2 and 3 of the PATH study, by relating initiation of cigarette smoking at Wave 3 to vaping at Wave 2, restricting attention to those who, at Wave 1, had never vaped, and using propensity indicators recorded at Wave 1 linked to uptake of e-cigarettes by Wave 2.

Here we describe the results of extensive analyses conducted based on Waves 1, 2 and 3 which not only include the main analyses envisaged at the time of our earlier paper [8], but also a variety of sensitivity and alternative analyses.

Methods

Some aspects of the analyses described here are the same as those described earlier [8] and are not presented again here. The selection of demographic and other predictor variables is the same as before, except that in some analyses we use exact age (12, 13, 14, 15, 16 and 17) which could now be estimated from the age group (12–14, 15–17) at the three Waves and the Wave when youths became adults (18+) for the first time. Use of the person-level weights provided in the PATH study database is as before, as is the process by which a sequence of logistic regression analyses is used to develop the shorter list of demographic variables to be used in forming the propensity scores.

Our main analysis M1 is based on those with data at Waves 1, 2 and 3 who had never smoked cigarettes by Wave 2 and had never used e-cigarettes by Wave 1. This analysis predicts Wave 3 ever smoking from Wave 2 ever e-product use, with adjustment based on Wave 1 predictors used to derive a propensity index for taking up e-products between Waves 1 and 2, and exact age being used in preference to grouped age. Note that, whereas in Wave 1 questions in PATH related only to e-cigarette use, in Waves 2 and 3 questions related to ever e-product use, which also included use of e-cigars, e-pipes and e-hookahs.

Associated with main analysis M1 are four sensitivity analyses (S1 to S4) which are otherwise similar, except that:

- S1. Those who had ever used other tobacco products at Wave 1 are excluded;
- S2. Ever use of other tobacco products at Wave 1 is included as an additional predictor variable;
- S3. The analysis is based on a propensity score for ever cigarette smoking rather than for ever vaping; or
- S4. The analysis is based on a propensity score for ever use of any tobacco product rather than for ever vaping.

Note that in our original paper [8] we also presented results of a further sensitivity analysis, based on linking current vaping to current smoking. This was not repeated here as numbers of new current smokers in current vapers were very low.

Main analysis M2 is similar to M1, except that analysis adjusts for the propensity index as originally derived [8], based on 12 variables recorded at Wave 1. Alternative versions of M2 substitute exact age rather than grouped age in deriving the propensity index, and/or included Wave 1 vapers in the analysis.

Main analysis M3 adjusts for a propensity index derived by linking Wave 2 predictors to Wave 2 e-product use. This is a replicate of the analysis conducted originally [8], but using a different period of taking up cigarettes. Data for Wave 1 were ignored, except that where the data for a characteristic was “ever in last 12 months”, Wave 1 data were used to define “ever”. An alternative version of M3 replaces grouped age by exact age in deriving the propensity index.

Apart from analyses linking Wave 2 e-product use to additional cigarette smoking at Wave 3 in those who had never smoked at Wave 2, two additional analyses (A1 and A2) were also conducted.

Additional analysis A1 relates e-cigarette use at Wave 1 to cigarette smoking at Wave 2 as in our earlier publication [8], but based on individuals who provided data at all three Waves. One version of this uses the same 12 variables as before to develop the propensity index, the other replaces grouped age by exact age. The OR from this analysis can be combined with that reported for main analysis M2 to give a combined estimate of the gateway effect for Wave 1 to 2 initiation and Wave 2 to 3 initiation based on the same set of variables determined at Wave 1.

Additional analysis A2 ignores Wave 2 data and relates e-cigarette use at Wave 1 to cigarette smoking at Wave 3 using the same 12 variables as before, but replacing grouped age by exact age.

Consideration of residual confounding was also taken into account for three of the analyses described above (M1, M3, A1), all involving exact age. In each case, the list of predictor variables was unaltered from that used originally, but the values of the predictor variables and of the propensity index were revised based on data available at all three Waves. For age, individual year of age at Wave 1 was used, while gender and Hispanic origin did not change between Waves. For the other variables used to form the propensity index, we used all the available data, generally choosing the response most associated with increased e-cigarette use where response varied between Waves (see Additional File Table 1 for further details).

Table 1
 Predicting Wave 2 ever-cigarette use from 16 Wave 1 predictor variables (Main analysis M1)

Variable ^a	Levels	N	OR (95% CI)
Exact age	12	1518	1.00 (base)
	13	1474	1.71 (1.23–2.38)
	14	1451	1.97 (1.43–2.71)
	15	1376	2.25 (1.65–3.08)
	16	1188	2.55 (1.86–3.51)
	17	1051	3.75 (2.72–5.15)
Ever been curious about smoking a cigarette			0.86 (0.76–0.97) ^b
Think you will smoke a cigarette in the next year			0.59 (0.48–0.71) ^c
Anyone who lives with you now use tobacco	Cigarettes, cigars, cigarillos, filtered cigars	2140	1.00 (base)
	Smokeless or other tobacco only	319	1.73 (1.26–2.37)
	No-one living in the home uses tobacco	5599	0.78 (0.65–0.94)
Ever used alcohol at all	Yes	2483	1.00 (base)
	No	5575	0.53 (0.45–0.62)
Agree/disagree: like new and exciting experiences, even if I have to break the rules	Strongly agree	285	1.00 (base)
	Agree	1252	0.71 (0.52–0.97)
	Neither agree nor disagree	2107	0.64 (0.47–0.87)
	Disagree	2404	0.38 (0.28–0.53)
Youth's grade performance in school in past 12 months	Strongly disagree	2010	0.46 (0.32–0.65)
	Mostly A's	2342	1.00 (base)
	A's or B's	2849	1.30 (1.07–1.58)
	Mostly B's	702	1.60 (1.22–2.10)
	B's or C's	1346	1.47 (1.17–1.85)
	Mostly C's	325	2.16 (1.52–3.09)

Note: The model is based on 8058 youths with data on all 16 predictors who neither smoked nor used e-cigarettes at Wave 1.

^a The variables are shown in order of their inclusion into the model

^b The OR is per unit of the graded variable which represents decreasing curiosity

^c The OR is per unit of the graded variable which represents decreasing likelihood

^d The OR is per unit of the graded variable which represents decreasing likelihood, with those originally entered as missing because they thought that they would not smoke a cigarette in the next year scored as "definitely not" (Level 4)

Variable ^a	Levels	N	OR (95% CI)
	C's or D's	334	2.74 (1.95–3.86)
	Mostly D's	45	2.09 (0.90–4.87)
	D's or F's	71	2.54 (1.34–4.81)
	Mostly F's	10	1.85 (0.26–12.91)
	School is ungraded	34	1.80 (0.54–6.06)
How often you visit your	Several times a day	2464	1.00 (base)
Facebook, Google Plus,	About once a day	2284	0.67 (0.56–0.80)
MySpace, Twitter or other	3–5 days a week	1006	0.73 (0.58–0.92)
	1–2 days a week	732	0.51 (0.37–0.69)
	Never	1572	0.40 (0.31–0.53)
Agree/disagree: I think I	Strongly agree	18	1.00 (base)
would enjoy using tobacco	Agree	95	0.42 (0.14–1.31)
	Disagree	1517	0.57 (0.20–1.58)
	Strongly disagree	6428	0.35 (0.12–1.01)
Hispanic origin	Hispanic	2332	1.00 (base)
	Not Hispanic	5726	0.67 (0.57–0.79)
Became very distressed when	Past month	1940	1.00 (base)
something reminded of past	2–12 months	1137	0.86 (0.70–1.07)
	Over a year	906	0.71 (0.55–0.92)
	Never	4075	0.74 (0.62–0.89)
Cigarettes or tobacco might be	Yes	1057	1.00 (base)
available to youth at parent	No	7001	0.65 (0.52–0.80)
or guardian's home			
Money received in total during	None	2771	1.00 (base)
an average week	Less than \$1	331	1.34 (0.91–1.96)
	\$1 to \$5	1234	1.26 (0.99–1.61)
	\$6 to \$10	1019	1.40 (1.10–1.79)
	\$11 to \$20	1289	1.42 (1.14–1.77)
	\$21 to \$50	751	1.36 (1.06–1.75)
	\$51 to \$100	337	1.53 (1.11–2.10)

Note: The model is based on 8058 youths with data on all 16 predictors who neither smoked nor used e-cigarettes at Wave 1.

^a The variables are shown in order of their inclusion into the model

^b The OR is per unit of the graded variable which represents decreasing curiosity

^c The OR is per unit of the graded variable which represents decreasing likelihood

^d The OR is per unit of the graded variable which represents decreasing likelihood, with those originally entered as missing because they thought that they would not smoke a cigarette in the next year scored as "definitely not" (Level 4)

Variable ^a	Levels	N	OR (95% CI)
	\$101 to \$150	160	2.02 (1.33–3.06)
	\$151 or more	166	1.96 (1.29–2.99)
Last time 2 + times: had a hard time paying attention at school, work or home	Past month	2700	1.00 (base)
	2–12 months	1402	0.75 (0.62–0.92)
	Over a year	819	0.84 (0.64–1.09)
	Never	3137	0.72 (0.59–0.87)
Number of times seen Movie 4	Never	6839	1.00 (base)
	Once	858	0.91 (0.73–1.11)
	Twice	190	1.24 (0.83–1.86)
	3 or more times	171	1.91 (1.29–2.82)
Think you will try a cigarette soon			1.99 (1.17–3.37) ^d
Note: The model is based on 8058 youths with data on all 16 predictors who neither smoked nor used e-cigarettes at Wave 1.			
^a The variables are shown in order of their inclusion into the model			
^b The OR is per unit of the graded variable which represents decreasing curiosity			
^c The OR is per unit of the graded variable which represents decreasing likelihood			
^d The OR is per unit of the graded variable which represents decreasing likelihood, with those originally entered as missing because they thought that they would not smoke a cigarette in the next year scored as “definitely not” (Level 4)			

For analyses M1, M3 and A1, alternative versions were also run in which the number of variables adjusted for was increased by also including interactions of age with each of the other three predictors most strongly linked to the relevant gateway effect.

Results

M1 Relating initiation of cigarette smoking between Waves 2 and 3 to ever e-product use at Wave 2, with adjustment for Wave 1 predictors linked to uptake of e-cigarettes between Waves 1 and 2

Initial analyses linked exact age, four other demographic variables (gender, Hispanic origin, race and census region) and 60 other selected predictor variables to ever e-product use at Wave 2 in those who had not smoked or used e-cigarettes at Wave 1. A propensity index based on 16 variables was derived using the three step process described earlier [8]. Additional File Table 2 shows the steps at which different variables were eliminated from consideration, while Table 1 gives the fitted equation for the propensity index.

Table 1 about here

As shown in Table 2 adjustment for propensity removed about half the excess OR (i.e. OR – 1), the unadjusted OR of 5.60 (95% CI 4.52–6.93) reducing to either 3.37 (2.65–4.28) or 3.11 (2.47–3.92), depending on whether adjustment was as a continuous variable or as quintiles. A similar reduction in the OR, to 3.27 (2.57–4.16), was achieved by adjusting for the 16 variables individually. It can also be seen that, for the first seven variables adjusted for, the adjusted OR decreased steadily, to 3.25. Further adjustment had little or no effect, with introducing additional variables sometimes slightly increasing the estimated OR and sometimes slightly decreasing it.

Table 2
Relating Wave 3 ever smoking to Wave 2 ever e-product use (Main analysis M1)

Adjustment variables	OR (95% CI)
None	5.60 (4.52–6.93)
Propensity score as quintiles	3.11 (2.47–3.92)
Propensity score as a continuous variable	3.37 (2.65–4.28)
Exact age	4.87 (3.91–6.06)
+Ever been curious about smoking a cigarette	4.27 (3.41–5.34)
+Think you will smoke a cigarette in the next year	3.84 (3.06–4.82)
+Anyone who lives with you now use tobacco	3.73 (2.97–4.69)
+Ever used alcohol at all	3.48 (2.76–4.38)
+Agree/disagree: Like new and exciting experiences even if I have to break the rules	3.39 (2.68–4.28)
+Youth's grade performance in school in past 12 months	3.25 (2.57–4.12)
+How often you visit your Facebook, Google Plus, MySpace, Twitter or other	3.17 (2.50–4.01)
+I think I would enjoy using tobacco	3.17 (2.50–4.02)
+Hispanic origin	3.22 (2.54–4.09)
+Last time a significant problem with: becoming very distressed when something reminded of past	3.19 (2.51–4.05)
+Cigarettes or tobacco might be available to youth at parent or guardian's home	3.17 (2.50–4.02)
+Money received in total during an average week	3.25 (2.56–4.13)
+Last time 2 + times: Had a hard time paying attention at school, work or home	3.22 (2.53–4.09)
+Number of times seen Movie 4	3.28 (2.57–4.17)
+Think you will try a cigarette soon	3.27 (2.57–4.16)
Notes: The table shows the effects of adjustment based on the Wave 1 predictors used to derive a propensity index for taking up e-products between Wave 1 and 2. The analyses are based on those with data at Waves 1, 2 and 3 who had never smoked cigarettes by Wave 2 and had never used e-cigarettes by Wave 1. Between Waves 2 and 3, 3261/7367 (3.54%) of never users of e-products at Wave 2 took up smoking, while 148/893 (16.57%) of ever users did so. For individuals who were 16–17 at Wave 1, adult data were used to determine e-product use and cigarette smoking at later Waves. The table includes the results of a stepwise regression based on successively including the most significant adjustment variables, given that ever e-product use at Wave 2 was included in the model.	

Four sensitivity analyses of M1 were carried out, fuller details being given in Tables 3 to 6 of the Additional File.

Compared to M1, S1 excluded those who had ever used products other than cigarettes or e-cigarettes at Wave 1, both in the construction of the propensity index and in estimating the gateway effect. Whereas M1 involved 8260 youths, of which 409 initiated smoking between Waves 2 and 3, S1 involved 7945, of which 359 took up smoking. The propensity index developed for S1 involved all the 16 variables shown in Table 2, except for “Number of times seen Movie 4” and “Think you will try a cigarette soon”. Here, the pattern of results is similar to that for Table 2, with the unadjusted OR of 5.66 (95% CI 4.49–7.13) reducing to either 3.45 (2.67–4.46), 3.24 (2.53–4.15), or 3.23 (2.49–4.18), depending on whether adjustment was made for propensity as a continuous variable, propensity as quintiles, or all the 14 variables individually.

Compared to M1, the only difference for S2 was that ever smoked other tobacco products at Wave 1 was added to the 16 variables used in M1 to make up the propensity score, and was forced into the regression models. Starting with the same unadjusted OR as M1, the adjusted ORs were very similar; 3.37 (2.64–4.29), 3.07 (2.44–3.87) and 3.20 (2.50–4.08), after adjustment for propensity (continuous), propensity (quintiles), or all the individual variables.

Whereas M1 (and S1 and S2) adjusted for variables found to be predictive of initiating e-product use at Wave 2, S3 adjusted for variables predictive of cigarette smoking. Here, the final model included 27 variables. The unadjusted OR of 5.65 (95% CI 4.55–7.01)

slightly differed from that in M1 as the individuals considered had to have non-missing data on 27 variables rather than 16. However, the overall effect of adjustment was again similar, with the OR reducing to 3.28 (2.56–4.22) after adjustment for all 27 variables. As for M1, adjustment for the first four variables had the most effect. Adjustment for the first seven variables reduced the OR to 3.26 (2.57–4.13), similar to the OR after adjustment for all 27. Propensity adjustment was not carried out in S3.

Compared to M1, S4 adjusted for variables predictive of take-up of any tobacco product between Waves 1 and 2. Here, the propensity index was based on 18 variables, with the unadjusted OR of 5.74 (4.55–7.23) reducing to 3.31 (95% CI 2.56–4.28), 3.19 (2.48–4.09), or 3.21 (2.47–4.18), after adjustment for propensity (continuous), propensity (quintiles), or all the individual variables. Adjustment for all 18 variables had a similar effect to adjustment for the most important 10 variables, where the OR was 3.20 (2.47–4.14).

M2 Relating initiation of cigarette smoking between Waves 2 and 3 to ever e-product use at Wave 2, with adjustment for the same Wave 1 predictors as previously reported [8]

Here, instead of deriving the Wave 1 predictors linked to uptake of e-cigarettes between Waves 1 and 2, analysis M2 uses the same set of Wave 1 predictors used in our earlier work [8], the results being shown in Table 3. Here, the unadjusted OR of 5.74 (95% CI 4.62–7.13) reduced to 3.54 (2.81–4.45) after adjustment for propensity as quintiles and to 3.45 (2.72–4.37) after adjusting for the individual variables. While adjustment here removed about half the excess OR, the reduction was less, to 4.53 (3.62–5.68) after adjustment for propensity as a continuous variable. The reductions were similar if exact age rather than age group was included in the list of variables. Here the unadjusted OR was reduced to 3.51 (2.79–4.41) after adjustment for propensity as quintiles, 4.59 (3.66–5.74) after adjustment for propensity as a continuous variable, and 3.39 (2.67–4.30) after adjustment for the individual variables.

Table 3
Relating Wave 3 ever smoking to Wave 2 ever e-product use (Main analysis M2)

Adjustment variables	OR (95% CI)
None	5.74 (4.62–7.13)
Propensity score as quintiles	3.54 (2.81–4.45)
Propensity score as continuous variable	4.53 (3.62–5.68)
Age range	5.20 (4.17–6.49)
+ Ever used alcohol at all	4.45 (3.54–5.58)
+ Ever been curious about smoking a cigarette	4.10 (3.26–5.16)
+ Think you will smoke a cigarette in the next year	3.70 (2.94–4.68)
+ Agree/disagree: Prefer friends who are exciting and unpredictable	3.65 (2.89–4.61)
+ Reaction if parent/guardian found you using tobacco	3.64 (2.88–4.60)
+ Gender	3.63 (2.87–4.58)
+ Agree/disagree: I think I would enjoy using tobacco	3.63 (2.87–4.59)
+ Agree/disagree: Some products are safer than others	3.63 (2.87–4.59)
+ Ever used prescription drug not prescribed to you: Ritalin or Adderall	3.67 (2.90–4.64)
+ Has a Facebook, Google Plus, MySpace, Twitter or other social networking	3.53 (2.79–4.47)
+ Anyone who lives with you now use tobacco	3.45 (2.72–4.37)
Notes: The table shows the effects of adjustment based on the same Wave 1 predictors as used in our original paper [8]. The analyses are based on those with data at Waves 1, 2 and 3 who had never smoked cigarettes by Wave 2 and had never used e-cigarettes by Wave 1. Between Waves 2 and 3, 249/7133 (3.49%) of never users of e-products at Wave 2 took up smoking, while 146/880 (16.59%) of ever users did so. For individuals who were 16–17 at Wave 1, adult data were used to determine e-product use and cigarette smoking at later Waves. The table includes the results of a stepwise regression based on successively including the most significant adjustment variables, given that ever e-product use at Wave 2 was included in the model.	

Similar analyses were also run which did not exclude those who had used e-cigarettes by Wave 1. This increased the number of ever e-product users who took up smoking from 146 to 201, and slightly increased the unadjusted OR to 5.95 (4.89–7.23). However, the pattern

of decline following adjustment was quite similar. For example, the OR adjusted for the individual variables reduced to 3.31 (2.65–4.12) using grouped age and to 3.26 (2.62–4.06) using exact age.

M3 Relating initiation of cigarette smoking between Waves 2 and 3 to ever e-product use at Wave 3, with adjustment for Wave 2 predictors

As noted in the Methods section, M3 is essentially a replicate of our earlier work [8], but using a different period of introduction of cigarettes. The propensity score developed was based on 18 variables, using age group or exact age as alternatives. The results, shown in Table 4, indicate that, as earlier [8], a large proportion of the unadjusted association can be explained by adjustment. The largest proportion was explained by adjusting for the 18 variables making up the propensity score, with the unadjusted OR of 6.70 (95% CI 5.40–8.32) reducing to 2.25 (1.74–2.91) or 2.75 (1.75–2.93) depending on whether the list of variables included age range or exact age. However, most of this reduction could be explained by adjustment for propensity.

Table 4
 Relating Wave 3 ever smoking to Wave 2 ever e-product use (Main analysis M3)

Adjustment variables	Using age group OR (95% CI)	Using exact age OR (95% CI)
None	6.70 (5.40–8.32)	6.70 (5.40–8.32)
Propensity score as quintiles	2.77 (2.19–3.50)	2.74 (2.17–3.48)
Propensity score as a continuous variable	2.57 (1.98–3.33)	2.60 (2.00–3.36)
Age range	5.78 (4.62–7.22)	-
Exact age	-	5.45 (4.36–6.83)
+Last time a significant problem with: feeling very trapped, lonely, sad, blue, depressed	5.22 (4.17–6.54)	4.95 (3.94–6.21)
+Reaction if parent/guardian found you using tobacco	4.89 (3.89–6.14)	4.66 (3.70–5.87)
+Money received in total during an average week	4.65 (3.69–5.86)	4.52 (3.59–5.71)
+Number of times seen Movie 3	4.31 (3.41–5.44)	4.20 (3.32–5.31)
+Number of times seen Movie 4	4.12 (3.25–5.21)	4.02 (3.18–5.10)
+Ever been curious about smoking a cigarette	3.45 (2.71–4.38)	3.36 (2.64–4.28)
+Think you will smoke a cigarette in the next year	2.89 (2.26–3.70)	2.86 (2.24–3.66)
+Ever used alcohol at all	2.63 (2.05–3.37)	2.63 (2.05–3.38)
+In past 12 months, youth's grade performance at school	2.51 (1.95–3.22)	2.51 (1.95–3.23)
+Agree/disagree: using tobacco would help me calm down when I am angry	2.43 (1.89–3.12)	2.43 (1.89–3.13)
+How often you visit your social media accounts	2.43 (1.88–3.12)	2.45 (1.90–3.15)
+Would smoke if one of your friends offered you one	2.37 (1.84–3.06)	2.39 (1.86–3.09)
+Anyone who lives with you now use tobacco	2.34 (1.81–3.02)	2.36 (1.83–3.04)
+Think you will try a cigarette soon	2.33 (1.81–3.01)	2.35 (1.82–3.03)
+Agree disagree: some tobacco products are safer than others	2.30 (1.78–2.97)	2.32 (1.79–2.99)
+Youth has a curfew or set time to be home on school nights	2.29 (1.77–2.95)	2.30 (1.78–2.98)
+Ever used prescription drug not prescribed to you: Ritalin or Adderall	2.25 (1.74–2.91)	2.27 (1.75–2.93)
Notes: The table shows the effects of adjustment based on Wave 2 predictors linked to use of e-products in Wave 2. The analyses are based on those with data at Waves 2 and 3 ignoring data from Wave 1. Between Waves 2 and 3, 228/8233 (2.77%) of never users of e-products at Wave 2 took up smoking, while 145/949 (15.28%) of ever users did so. For individuals who were 17 at Wave 2, adult data were used to determine cigarette smoking at Wave 3. The table includes the results of a stepwise regression based on successively including the most significant adjustment variables, given that ever e-product use at Wave 2 was included in the model. The first set of ORs is based on a model including age group, while the second is based on a model including exact age.		

Combining the Wave 2 to 3 results shown in Table 4 with the Wave 1 to 2 results reported earlier [8] by fixed-effect meta-analysis, gives an unadjusted OR of 6.30 (5.31–7.46) which is reduced to 2.65 (2.24–3.18), 2.53 (2.07–3.10) or 2.08 (1.70–2.54) depending on whether adjustment is for propensity (quintiles), propensity (continuous) or all the variables making up the propensity score. This represents reductions in the excess OR of, respectively, 68.9%, 71.1% or 79.8%.

A1 Relating initiation of cigarette smoking between Waves 1 and 2 to ever e-cigarette use at Wave 1, based on individuals who provided data at all three Waves

Table 5 summarizes the main results of these analyses and compares them with those reported earlier [8]. While the original analyses were based on 9423 youths, 421 of whom initiated smoking, the new analyses were based on 8700 youths, 389 of whom initiated smoking. As can be seen, the results in the original analysis, based on grouped age, were similar to those from the new analyses, whether grouped or exact age was used.

Table 5
Relating Wave 2 ever smoking to Wave 1 ever e-cigarette use - original [8] and A1 ORs

Adjustment variables	Data on two Waves		Data on all three Waves
	Originally reported	Grouped age	Exact age
	OR (95% CI)	OR (95% CI)	OR (95% CI)
None	5.70 (4.33–7.50)	5.99 (4.52–7.95)	5.99 (4.52–7.95)
Propensity score as quintiles	2.48 (1.85–3.31)	2.65 (1.96–3.58)	2.59 (1.92–3.50)
Propensity score as continuous variable	2.47 (1.79–3.42)	2.67 (1.92–3.72)	2.64 (1.89–3.68)
Grouped age	4.81 (3.64–6.35)	5.04 (3.78–6.72)	-
Exact age	-	-	4.81 (3.60–6.42)
+ 11 further variables	1.85 (1.35–2.53)	1.97 (1.42–2.73)	1.98 (1.43–2.75)

Notes: Each set of ORs is based on those who had never smoked cigarettes by Wave 1. The first analysis is as summarized in Table 1. The last two analyses only exclude those without data at Wave 3.

The results from analysis A1 for grouped age may theoretically be combined with those from analysis M2 shown in Table 3, as they both use the Wave 1 predictors from our original paper [8], with exact age replacing grouped age, and are both based on individuals with data at all three Waves. However, as illustrated by the results adjusted for all 12 variables, where the ORs are 3.45 (95% CI 2.72–4.37) from Tables 3 and 1.97 (1.42–2.73) from Table 5, these estimates are heterogeneous ($p < 0.001$), providing a random-effects combined estimate of 2.64 (1.52–4.57).

A2 Relating Wave 3 ever smoking to Wave 1 e-cigarette use, ignoring Wave 2 data

This analysis is similar to that reported originally [8] but relates to a longer follow-up period, and uses exact rather than grouped age. The results of this analysis, shown in Table 6, are quite similar to those shown in Table 5. Again, an unadjusted OR is markedly reduced by adjusting for propensity, whether as quintiles or as a continuous variable, and is further reduced by adjusting for all the 12 individual variables considered.

Table 6
Relating Wave 3 ever smoking to Wave 1 ever e-cigarette use using exact age

Adjustment variables	OR (95% CI)
None	5.65 (4.50–7.10)
Propensity score as quintiles	2.48 (1.95–3.16)
Propensity score as continuous variable	2.61 (2.00–3.40)
Exact age	4.69 (3.71–5.93)
+ 11 further variables	1.97 (1.51–2.56)

Notes: The table shows the effects of adjustment based on the same Wave 1 predictors as used in our original paper [8] but replacing age range by exact age. The set of ORs is based on those with data at Waves 1, 2 and 3 who had never smoked cigarettes by Wave 1. Between Waves 1 and 3, 716/8334 (8.59%) of never users of e-cigarettes at Wave 1 took up smoking, while 123/366 (33.61%) of ever users did so. The table includes the results of a stepwise regression based on successively including the most significant adjustment variables, given that ever e-product use at Wave 1 was included in the model.

Attempting to account for residual confounding

Table 7 summarizes the main results shown in Table 2 for main analysis M1, which make no allowance for residual confounding, and compares them with the results of an analysis using the same list of predictor variables, but with values modified in an attempt to adjust for residual confounding. As can be seen, markedly more of the unadjusted association was explained when allowance for

residual confounding was made, with the adjusted ORs in the range 2.36 to 2.46 when allowance was made, compared with 3.11 to 3.37 when it was not. Note that the unadjusted ORs in the two sets of results vary slightly, as missing values in some individuals in the original analyses were replaced by estimates taken from other Waves.

Table 7
Effect of allowance for residual confounding in main analysis M1

Adjustment variables	M1 – no allowance	M1 – allowance
	OR (95% CI)	OR (95% CI)
None	5.60 (4.52–6.93)	5.65 (4.58–6.98)
Propensity score as quintiles	3.11 (2.47–3.92)	2.40 (1.91–3.02)
Propensity score as a continuous variable	3.37 (2.65–4.28)	2.46 (1.93–3.14)
All 16 variables individually	3.27 (2.57–4.16)	2.36 (1.85–3.02)
Notes: The “no allowance” results correspond to those in Table 6.		

The analyses are based on those with data at Waves 1, 2 and 3 who had never smoked cigarettes by Wave 2 and had never used e-cigarettes by Wave 1. Between Waves 2 and 3 261/7367 (3.54%) of never users of e-products at Wave 2 took up smoking, while 148/893 (16.57%) of ever users did so in the population considered in the “no allowance” analyses. The corresponding figures in the “allowance” analyses were 267/7682 (3.48%) and 150/915 (16.39%). For individuals who were 16–17 at Wave 1, adult data were used to determine e-product use and cigarette smoking at later Waves. The table includes the results of a stepwise regression based on successively including the most significant adjustment variables, given that ever e-product use at Wave 2 was included in the model.

While allowance for residual confounding has quite a marked effect for analysis M1, the analysis which avoided the possibility of over-adjustment, it did not for analyses M3 and A2, which did not avoid this possibility. Detailed results are shown in Tables 7 and 8 in the Additional File.

Table 8
Summary of results from analyses

Analysis	Baseline		Follow-up		Predictor	Age	Comment	Unadjusted OR	% Excess OR explained ^a			
	Wave	Wave	Wave	Wave					P as Q ^b	P as C ^c	6 variables	All variables
A	Original	1	2	2	Ever e-cigs	Grouped	As published ([8])	5.70	68.5	68.7	78.1	81.9
B	M1	2	3	3	Ever e-cigs	Exact	Predictors revised based on those who were not Wave 1 e-users	5.60	54.1	48.5	48.0	50.7
C	M1/S1	2	3	3	Ever e-cigs	Exact	As M1 but excludes Wave 1 other product users	5.66	51.9	47.4	47.9	52.1
D	M1/S2	2	3	3	Ever e-cigs	Exact	As M1 but Wave 1 other product use included as predictor	5.60	55.0	48.3	50.2	52.2
E	M1/S3	2	3	3	Ever e-cigs	Exact	As M1 but adjusting for predictors of ever cigarette smoking	5.65	-	-	48.4	51.0
F	M1/S4	2	3	3	Ever any product	Exact	As M1 but adjusting for predictors of ever any tobacco use	5.74	53.8	51.3	45.1	53.4
G	M2	2	3	3	Ever e-cigs	Grouped	Original 12 predictors	5.74	46.4	25.5	44.3	48.3
H	M2 (variant)						Did not exclude Wave 1 e-users	5.95	50.3	28.7	49.3	53.3
I	M2 (variant)	2	3	3	Ever e-cigs	Exact	Original 12 predictors	5.74	47.0	24.3	46.0	49.6
J	M2 (variant)						Did not exclude Wave 1 e-users	5.95	50.7	28.3	50.5	54.3

^a % excess explained = $100 \cdot (OR_U - OR_A) / (OR_U - 1)$ where OR_U is the unadjusted OR, and OR_A is the adjusted OR

^b P as Q = Propensity as quintiles

^c P as C = Propensity as a continuous variable

		Baseline	Follow-up				Unadjusted	% Excess OR explained ^a			
K	M3	2	3	Ever e-cigs	Grouped	Predictors revised essentially ignoring Wave 1 data	6.70	68.9	72.5	45.3	78.1
L	M3 (variant)				Exact	As above but using exact age	6.70	69.5	71.9	47.0	77.8
M	A1	1	2	Ever e-cigs	Grouped	As original but based on those with data on all three Waves	5.99	66.9	66.5	76.8	80.6
N	A1 (variant)				Exact	As above but using exact age	5.99	68.1	67.1	77.0	80.4
O	A2	1	3	Ever e-cigs	Exact	Original predictors but ignoring Wave 2	5.65	68.2	65.4	74.4	79.1
P	M1 (variant)	2	3	Ever e-cigs	Exact	As M1 but allows for residual confounding	5.65	69.9	68.6	60.0	70.8
Q	M3 (variant)	2	3	Ever e-cigs	Exact	As M3 but allows for residual confounding	6.67	75.3	74.3	51.5	80.2
R	A1 (variant)	1	2	Ever e-cigs	Exact	As A1 but allows for residual confounding	6.10	69.0	68.0	65.1	76.7
^a % excess explained = $100 \times (OR_U - OR_A) / (OR_U - 1)$ where OR_U is the unadjusted OR, and OR_A is the adjusted OR											
^b P as Q = Propensity as quintiles											
^c P as C = Propensity as a continuous variable											

Investigating whether introducing some interactions explains more of the gateway effect

Versions of analyses M1, M3 and A1 were also seen, in which the number of variables adjusted for was extended by also including interactions of age with each of the other three predictors most strongly linked to the gateway effect. For analysis M1, allowance for these interactions had virtually no effect, the original estimate of 3.27 (95% CI 2.57–4.16) shown in Table 2 without including interactions changing to 3.26 (2.55–4.15) when interactions were included in the model. For analysis M3, the estimate changed only from 2.27 (1.75–2.93) to 2.35 (1.81–3.05), while for analysis A1, it changed from 1.98 (1.43–2.75) to 2.06 (1.48–2.88).

Summary of results

Table 8 summarizes the results from 18 of the analyses described above, expressing the extent to which adjustment explained the unadjusted OR using the statistic $100 \times (OR_U - OR_A) / (OR_U - 1)$ where OR_U is the unadjusted OR, and OR_A is the adjusted OR. The most obvious impression from the table is that the results largely fall into two groups.

Table 8 about here

Results from the original analysis and for analyses M3, A1 and A2 (rows A, K to O, and Q to R of Table 8) all show that as much as about 80% of the unadjusted excess OR can be explained by adjustment for the full set of variables in the model, with somewhat less, typically about 70%, explained using propensity as quintiles or as a continuous variable.

In contrast, results from virtually all of analyses M1 and M2 (rows B to K) show that only about 50% of the unadjusted excess OR can be explained by adjustment for the full set of variables, with propensity as quintiles giving generally similar results.

The difference between these two groups is that the first set of results are subject to the problem of over-adjustment, with the values of the predictors used possibly having been affected by having used e-cigarettes. This is mainly so where the baseline wave was Wave 1, but was also true for analysis M3 where Wave 1 data were essentially ignored. In contrast the second set of results avoided over-adjustment by considering follow-up from Wave 2 to 3, with predictors based on Wave 1 data in youths who had never used e-cigarettes. However in this second set of results the variables used were not as up-to-date as in the first analyses.

The variant analysis of M1, allowing for residual confounding (row P) gives an intermediate result with about 70% of the excess risk being explained, whether by the full set of variables or by propensity. This analysis, however, does not avoid the problem of over-adjustment as it incorporates some information from waves where individuals were already using e-cigarettes.

It is clear from Table 8 that many of the variables studied had little effect on the pattern of results. These included use of grouped or exact age, taking into account use of other products, and using predictors of cigarette smoking or any tobacco use rather than predictors of e-cigarette use.

Two other conclusions may be drawn from Table 8. One is that adjustment for propensity as quintiles or as a continuous variable generally gives very similar results, with the exception of analysis M2 and its variants, where propensity as a continuous variable explained substantially less of the unadjusted excess OR. Inspection of the detailed modelling results showed that, whereas in other analyses, the logarithm of the OR increased fairly linearly with quintiles of propensity, in the case of analysis M2 and its variants it did not. Thus, in M1 for example, the log ORs by quintile were 0, 0.73, 1.11, 1.66 and 2.52 while in M2 they were 0, 0.21, 0.96, 1.51 and 2.19 with very little rise between quintiles 1 and 2.

The other is that adjustment for the first six variables in the model generally explained a very substantial part of the unadjusted excess OR explained by the full set. Though this was not true for analysis M2, it was still true that adjustment for the last eight or nine variables explained far less of the excess OR than did the first eight or nine.

Discussion

In our publication based on Waves 1 and 2 [8] our analyses showed that an unadjusted estimate of the gateway effect 5.70 (85% CI 4.33–7.50) could be considerably reduced by adjustment, to 1.59 (1.14–2.20) in the most striking case. Because of the marked reduction in the OR following adjustment, and the possibility of incomplete control for confounding we regarded it as “unclear whether prior vaping actually increases uptake of cigarette smoking”. However, we did note the possibility of over-adjustment, with vaping at Wave 1 possibly having affected the recorded values of some of the variables used for adjustment.

At that time we noted that this possibility of over-adjustment could be addressed in analyses relating initiation of cigarette smoking at Wave 3 to vaping at Wave 2, restricting attention to those youths who, at Wave 1, had never vaped, and using adjustment variables recorded at Wave 1. This we have done in the analyses reported here, and our major finding is that adjustment reduced the excess risk far less, by only about 50% rather than about 80%, in our main analysis M1.

While these results more strongly support the existence of a true gateway effect of taking up vaping, there must still remain doubt about its magnitude. One reason is that predictors recorded a year before the baseline may not fully account for the characteristics of the youth at the start of follow-up. A second reason is that, although the PATH study records data on a whole range of possibly relevant characteristics, there may be some relevant predictors not considered. A third reason is that the answers to some of the questions may have been inaccurately measured. We have attempted to address this problem of residual confounding by amending values of predictors recorded at Wave 1 to take into account data recorded at later Waves. However, this problem re-introduces the problem of over-adjustment as Wave 2 and 3 values may have been affected by vaping. Theoretically, one could use data from Waves 1 to 4, using data for Waves 1 and 2 from youths who have never vaped to produce more accurate estimates of the predictors to use for a study of gateway effects between Waves 3 and 4. But this would add to the problem of using predictors recorded some time before follow-up.

Since the time that we published our earlier analysis [8] and our paper on general considerations relating to vaping as a possible gateway into cigarette smoking [7] a number of other authors have presented evidence from prospective studies [9–15]. The studies vary in the extent to which potential confounding variables have been adjusted for, with large OR estimates tending to be reported in studies with more limited control. Thus, a study in the Netherlands [15], which adjusted only for sex, age education and a single indicator of propensity to smoke, reported an OR of 11.90 (95% CI 3.36–42.11) for the relationship between ever use of e-cigarettes with nicotine and initiation of cigarette smoking during follow-up. Also, a study in the US [9], which adjusted only for demographic variables and use of other tobacco products, reported ORs of 7.08 (2.34–21.42) and 3.87 (1.86–2.06) depending on the follow-up period studied, while another US study [13], with limited control for confounding variables, reported an OR of 3.57 (1.96–6.45). Apart from a US study [14], which reported an OR of 6.8 (1.7–28.3), following adjustment for ten covariates independently associated with initiation of smoking, most of the other studies which appear to have better control for confounding gave lower estimates. These included a study in Taiwan [10] which reported an OR of 2.14 (1.66–2.75), a study in Germany [12] which reported an OR of 2.18 (1.65–2.87) and a study in Finland [11] which reported that adjustment reduced the OR from 11.52 (4.91–26.56) to 2.92 (1.09–7.85). Notably, a study in Great Britain [16] reported an OR of 11.89 (3.56–39.72) estimated using the usual logistic method, but a reduced value of 1.34 (1.05–1.72) using causal mediation analysis.

Generally our results are consistent with the literature in confirming that a substantial proportion, but not all, of the observed association between e-cigarette use and subsequent initiation of cigarette smoking can be explained by adjustment for factors linked to susceptibility to tobacco. However, large cohort studies with high quality, accurate, data on a wide range of predictive factors recorded at regular intervals will be needed to gain better insight into the magnitude of any true causal effect of vaping. The PATH study with its multiple waves and comprehensive questionnaire should prove more and more useful in the future.

There are a number of theoretical beneficial and adverse effects of e-cigarettes [7]. Beneficial effects relate to individuals who would otherwise have smoked vaping instead, smokers who would otherwise have continued to smoke switching instead to vaping, vaping helping established smokers to quit, and vaping helping established smokers to reduce their cigarette consumption. Apart from vaping encouraging initiation of smoking, theoretical adverse effects occur if smokers intending to quit switch to vaping instead, or smokers add vaping to their normal cigarette consumption. It is clearly important, therefore, to take all these considerations into account when attempting to estimate the health impact of e-cigarettes.

Conclusions

By using data from three Waves of the PATH study, the analyses of the gateway effect reported here improve on those reported earlier [8] based on the first two Waves by allowing potential confounding variables to be determined at a time before vaping started. Whereas the earlier analyses suggested that the adjustment for confounding explained about 80% of the unadjusted relationship between vaping and subsequent initiation of smoking, our current analyses suggest that adjustment explains only about 50%. This provides stronger evidence of a true effect of vaping, although doubt still remains about its true magnitude for reasons discussed.

Abbreviations

CI
confidence interval;
OR
odds ratio;
PATH
Population Assessment of Tobacco and Health

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The data from the PATH study are publicly available at <https://pathstudyinfo.nih.gov>

Competing interests

Both authors are long term consultants to the tobacco industry

Funding

Financial support was provided by Philip Morris Products SA, through Project Agreement No. 29 with P.N. Lee Statistics and Computing Ltd. While some technical comments were provided by the funder on drafts of the statistical plan and this publication, the final versions remain the responsibility of the authors.

Authors' contributions

PNL drafted the statistical plan, checked the analyses conducted, and drafted this publication. JSF commented on the plan, carried out the analyses, and commented on the publication. Both authors agreed the final publication.

Acknowledgements

We thank Esther Afolalu for assistance in acquiring the data from the PATH study, and Zheng Sponsiello-Wang and Christelle Chrea for providing technical comments at various stages. We also thank Jan Hamling for assistance in running the analyses, and Yvonne Cooper and Diana Morris for typing the various drafts of the paper.

References

1. Barrington-Trimis JL, Urman R, Leventhal AM, Gauderman WJ, Cruz TB, Gilreath TD, et al. E-cigarettes, cigarettes, and the prevalence of adolescent tobacco use. *Pediatrics* 2016;138(2).(Epub 20160711): doi:10.1542/peds.2015-3983.
2. Best C, van der Sluijs W, Haseen F, Eadie D, Stead M, MacKintosh AM, et al. Does exposure to cigarette brands increase the likelihood of adolescent e-cigarette use? A cross-sectional study. *BMJ Open* 2016;6(2):e008734.(Epub 20160223): doi:10.1136/bmjopen-2015-008734.
3. Miech R, Johnston L, O'Malley PM, Bachman JG, Patrick ME. Trends in adolescent vaping, 2017-2019 (Letter). *N Engl J Med* 2019;381(15):1490-1.(Epub 20190918): doi:10.1056/NEJMc1910739.
4. National Academies of Sciences Engineering and Medicine. Public health consequences of e-cigarettes. Washington DC: The National Academies Press; 2018. Available: <https://doi.org/10.17226/24952>.

5. Nutt DJ, Phillips LD, Balfour D, Curran HV, Dockrell M, Foulds J, et al. Estimating the harms of nicotine-containing products using the MCDA approach. *Eur Addict Res* 2014;20(5):218-25. doi:10.1159/000360220.
6. Soneji S, Barrington-Trimis JL, Wills TA, Leventhal AM, Unger JB, Gibson LA, et al. Association between initial use of e-cigarettes and subsequent cigarette smoking among adolescents and young adults: A systematic review and meta-analysis. *JAMA Pediatr* 2017;171(8):788-97.(Epub 20170626): doi:10.1001/jamapediatrics.2017.1488.
7. Lee PN, Coombs KJ, Afolalu EF. Considerations related to vaping as a possible gateway into cigarette smoking: an analytical review [version 3; peer review: 2 approved]. *F1000Research* 2018;7(1915):(https://doi.org/10.12688/f1000research.6928.3). doi:doi.org/10.12688/f1000research.16928.3.
8. Lee PN, Fry J. Investigating gateway effects using the PATH study [version 1; peer review: 1 approved]. *F1000Research* 2019;8(264):(https://doi.org/10.12688/f1000research.8354.1). doi:10.12688/f1000research.18354.1.
9. Bold KW, Kong G, Camenga DR, Simon P, Cavallo DA, Morean ME, et al. Trajectories of e-cigarette and conventional cigarette use among youth. *Pediatrics* 2018;141(1).(Epub 20171204): doi:10.1542/peds.2017-1832.
10. Chien YN, Gao W, Sanna M, Chen PL, Chen YH, Glantz S, et al. Electronic cigarette use and smoking initiation in Taiwan: Evidence from the First Prospective Study in Asia. *Int J Environ Res Public Health* 2019;16(7).(Epub 20190330): doi:10.3390/ijerph16071145.
11. Kinnunen JM, Ollila H, Minkkinen J, Lindfors PL, Timberlake DS, Rimpela AH. Nicotine matters in predicting subsequent smoking after e-cigarette experimentation: A longitudinal study among Finnish adolescents. *Drug Alcohol Depend* 2019;201:182-7.(Epub 20190619): doi:10.1016/j.drugalcdep.2019.04.019.
12. Morgenstern M, Nies A, Goecke M, Hanewinkel R. E-Cigarettes and the use of conventional cigarettes. *Dtsch Arztebl Int* 2018;115(14):243-8.
13. Pénczes M, Foley KL, Nadasan V, Paulik E, Ábrám Z, Urbán R. Bidirectional associations of e-cigarette, conventional cigarette and waterpipe experimentation among adolescents: A cross-lagged model. *Addict Behav* 2018;80:59-64.(Epub 20180111): doi:10.1016/j.addbeh.2018.01.010.
14. Primack BA, Shensa A, Sidani JE, Hoffman BL, Soneji S, Sargent JD, et al. Initiation of traditional cigarette smoking after electronic cigarette use among tobacco-naive US young adults. *Am J Med* 2018;131(4):443.e1-.e9.(Epub 20171211): doi:10.1016/j.amjmed.2017.11.005.
15. Treur JL, Rozema AD, Mathijssen JJP, van Oers H, Vink JM. E-cigarette and waterpipe use in two adolescent cohorts: cross-sectional and longitudinal associations with conventional cigarette smoking. *Eur J Epidemiol* 2018;33(3):323-34.(Epub 20171219):

doi:10.1007/s10654-017-0345-9.

16. East K, Hitchman SC, Bakolis I, Williams S, Cheeseman H, Arnott D, et al. The association between smoking and electronic cigarette use in a cohort of young people. *J Adolesc Health* 2018;62(5):539-47.(Epub 20180221): doi:10.1016/j.jadohealth.2017.11.301.

Supplementary Files

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

- [GatewaypaperforBMCPublicHealthAdditionalfile2.docx](#)