

# Unpacking the gateway hypothesis of e-cigarette use:

## The need for triangulation of individual- and population-level data

Lion Shahab, Ph.D<sup>1,2</sup>, Jamie Brown, Ph.D<sup>1,2</sup>, Lies Boelen, PhD<sup>3</sup>, Emma Beard, Ph.D<sup>1,2</sup>, Robert West, PhD<sup>1</sup>, Marcus R  
Munafò, PhD<sup>2,4,5</sup>

1 Department of Behavioural Science and Health, University College London, London, UK

2 SPECTRUM Consortium, Edinburgh, UK

3 Sandtable Ltd, London, UK

4 School of Psychological Science, University of Bristol, Bristol, UK

5 MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK

\*Correspondence to: Professor Lion Shahab, Department of Behavioural Science and Health, University College London, 1–19 Torrington Place, London, WC1E 6BT, UK; Phone: +44 207679 1895; Fax: +44 207813 2848; Email: lion.shahab@ucl.ac.uk

Accepted Manuscript

## Commentary

As recent publications in *N&TR* highlight,<sup>1</sup> there is considerable debate whether e-cigarettes act as a causal gateway to subsequent smoking in adolescents. Answering this question requires a clear specification of what we mean by “gateway” in this context. Unlike the original hypothesis, which postulated a causal progression from legal/“soft” (e.g., alcohol) to illicit/“hard” (e.g., heroin) polydrug use, initiated in adolescence,<sup>2</sup> current gateway discussions in tobacco control centre around the idea of what is essentially a transition from less (e-cigarette) to more harmful (cigarette) modalities of nicotine use.

This complicates causal inference; for the original hypothesis it was possible to assess changes in sensitivity to, and rewarding effects of, a destination drug following exposure to a gateway drug in animal models, but such highly controlled experimental work in model systems is impossible when both the putative gateway and destination drug are the same. Randomised controlled trials are equally unfeasible: it would be unethical to randomise adolescents to try or not try e-cigarettes and impractical to assess relatively long-term outcomes (i.e., transition to smoking, potentially several years later). We therefore must rely on observational data. Here we introduce the idea of a triangulation framework, using evidence from different approaches to provide robust answers to causal questions, and apply this to the problem of the gateway to assess the current balance of evidence.

Triangulation attempts to support stronger causal inference by considering findings across multiple methodological approaches, each with different sources of bias.<sup>3</sup> For research into gateway effects of e-cigarette use, we define two major categories of approach: i) individual-level approaches (i.e., where individual participants provide data), and ii) population-level approaches (i.e., using summary data from whole populations). Cutting across these, one can also distinguish between cross sectional comparisons between individuals or populations, and longitudinal comparisons of individuals or populations over time.

At the individual level, numerous cross-sectional and prospective studies show a strong positive association between e-cigarette use and smoking cigarettes.<sup>4</sup> Young people who report using e-cigarettes are more likely to report smoking, both concurrently and in the future. However, these studies are potentially subject to measurement error (e.g., misreporting of smoking status, especially among youth<sup>5</sup>) and confounding.<sup>6</sup> Just because e-cigarette use

precedes cigarette use does not mean that e-cigarette use *caused* subsequent smoking: adolescents that try e-cigarettes may have tried cigarettes anyway due to some underlying common liability, such as a genetic predisposition to risky behaviour.<sup>7</sup> Although it is possible to reduce such confounding statistically (e.g., propensity score matching),<sup>8</sup> success depends on all relevant confounders being included (which is unlikely, resulting in unmeasured confounding), included confounders being measured accurately (again unlikely, resulting in residual confounding), and the association with the outcome being modelled through an appropriate function (e.g. exponential). Even small model misspecifications can result in spurious associations.<sup>6</sup> No matter how many studies one conducts with the same sources of bias, the limitations of this approach remain, running the unintentional risk of creating the appearance of robustness and acceptance that causal claims are true.<sup>9</sup>

What is needed then are complementary and independent lines of evidence that address this problem of confounding by combining methodologies that have distinct potential biases. If results across approaches align, this provides greater confidence in a causal interpretation, since it is unlikely that different sources of bias would conspire to give the same result in each case. There are numerous methods that can be applied to individual-level data (Supplementary Table 1).<sup>3</sup> One of these is the use of ‘negative control’ outcomes; these exhibit a similar confounding structure to smoking, but have no plausible mechanism through which e-cigarette use could influence them (e.g. unprotected sexual intercourse). If e-cigarette use shows a similar association with this as with smoking, it suggests that confounding (e.g., via risk taking personality) is the most likely culprit rather than there being a true causal association.<sup>10</sup> Indeed, there are studies showing associations with such outcomes, including illicit prescription medication use.<sup>11</sup>

Instrumental variable analysis constitutes another individual-level approach. This would require an instrument causally related to e-cigarette use but without any plausible causal connection with smoking. Such studies could exploit environmental instruments or use genetic variants as proxies for the exposure (as in Mendelian Randomisation, MR).<sup>12</sup> One environmental instrument could be access to e-cigarettes, for example by density of vape shops in the locality, opportunity for online purchasing, or pricing of e-cigarettes. Thus, if it turns out that adolescents with greater access specifically to e-cigarettes are more likely to smoke, this would support a causal role for e-cigarettes in the take-up of smoking. Identifying genetic instruments that differentiate e-cigarette from

cigarette use may be more difficult. Early evidence indicates a common genetic vulnerability to both smoking and e-cigarette use, which may reflect a broad risk taking phenotype (in itself suggestive that at least part of the association between e-cigarette use and smoking may be non-causal).<sup>13</sup> Other genetic approaches beyond MR, including twin and sibship comparison studies,<sup>14</sup> may hold promise for disentangling genetic confounders and intergenerational transmission to assess true causal associations between e-cigarette use and smoking.

Turning to the population level (Supplementary Table 1), time series analyses can examine associations over time between prevalence of e-cigarette and cigarette use in potentially vulnerable age groups. Because the whole population is being studied, this rules out individual factors accounting for any association found. Population-level studies can be biased by confounding variables that operate at the population level, but these will not overlap with individual confounders. For example, a country might relax or tighten regulations for e-cigarettes and cigarettes at the same time or fund a media campaign targeting both types of product. Therefore, it remains important to adjust for population-level confounding as far as possible (e.g., shifts in policy). Population-level time series analyses have the additional advantage of directly estimating the population-level effect of e-cigarettes as has already been done for the impact of e-cigarette use on smoking cessation.<sup>15</sup> Importantly, such analysis should include appropriately lagged effects to assess any putative gateway.

Other methodological approaches at population-level include cross-context comparisons and natural experiments, which can exploit the likelihood that confounding structures will be different across populations in dissimilar contexts (either historically, e.g. because of differently patterned behaviours, or by design, e.g. due to legislative changes introduced in one but not in another context).<sup>3</sup> Similar findings across contexts could not be explained readily by confounding, while different results would be unlikely due to true differences in causal effects between populations. No studies of this kind have been reported. A preliminary look at population data shows that over the same time-period that e-cigarette use increased, cigarette use decreased, and this is found in markedly different contexts,<sup>16,17</sup> suggesting a limited role of confounding.

One can also use modelling (Supplementary Table 1) to assess the extent to which individual- and population-level estimates of gateway effect agree with each other, e.g. with micro-simulation or agent-based models, employed for now-or forecasting in economics, and increasingly applied to guide health policy.<sup>18</sup> Our group has developed such a model in collaboration with Sandtable, a data analysis company. Here, adolescent cigarette uptake is calibrated to match rates of decline prior to e-cigarettes becoming popular, e-cigarette uptake is matched to observed values and the model run to estimate smoking prevalence in counterfactual scenarios, using different associations between e-cigarette and cigarette use. We find that the gateway effect estimated from individual-level studies<sup>19</sup> predicted far less of a reduction in population smoking prevalence in the relevant age group than was actually observed (Figure 1), making it unlikely that individual-level effects were genuine.

Lastly, while we have focused on the analysis of individual- and population-level aggregated data, with the proliferation of mobile phone technologies, n-of-1 studies may also offer an insight into possible gateway effects (e.g., by following transitions of individuals in different contexts using ecological momentary assessment). Further, most designs and statistical methods mentioned here rely heavily on Frequentist hypothesis testing, which can lead to non-significant findings being conflated with evidence of no effects. One remedy is the use of Bayes factors which assess the extent of evidence for the null hypothesis and can also determine if the data analysis is insensitive or underpowered.<sup>20</sup> Many of the approaches mentioned in Supplementary Table 1 have Bayesian equivalents, which in themselves can offer a method of triangulation, as prior knowledge from already published studies can be incorporated.

If we carry on as we are, we are unlikely to address this important scientific and public health question satisfactorily and may merely perpetuate disagreement by selective reporting of results that favour one direction or the other of a gateway effect. In addition to encouraging single method investigations using new and diverse methodologies, we would therefore argue for a common, transparent approach that involves prospective registration of a triangulation framework based on individual and population-level data and *a priori* specification of what would be considered sufficient evidence in each direction of the postulated gateway, based not only on statistical significance but also on clinically or theoretically meaningful effect sizes (for an example see

<https://osf.io/nd2qk>). Taking a collaborative approach, involving groups with different prior beliefs, would not only increase feasibility of triangulation but also help unify our field. Based on the current balance of evidence, using triangulated data from recent population-level cross-contextual comparisons, individual-level genetic analyses and modelling, we do believe, however, that causal claims about a strong gateway effect from e-cigarettes to smoking are unlikely to hold, while it remains too early to preclude other smaller or opposing effects.

Accepted Manuscript

## ADDITIONAL INFORMATION

### Acknowledgments

None.

### Author contributions statement

LS and RW conceived of the original idea for this editorial and LS wrote the first draft. LB developed the microsimulation model. All authors read, reviewed and approved the final version. All researchers listed as authors are independent from the funders and all final decisions about the research were taken without constraint by the investigators. LS had final responsibility for the decision to submit for publication.

### Funding

None. The salary of LS is HEFCE funded. JB's and EB's salaries are funded by Cancer Research UK (C1417/A22962). LB is an employee of Sandtable, Ltd. LS, EB and JB are members of the SPECTRUM, a UK Prevention Research Partnership Consortium (MR/S037519/1). UKPRP is an initiative funded by the UK Research and Innovation Councils, the Department of Health and Social Care (England) and the UK devolved administrations, and leading health research charities.

### Conflict of Interest

LS has received a research grant and honoraria for a talk and travel expenses from manufacturers of smoking cessation medications (Pfizer and Johnson & Johnson). JB and EB have received unrestricted research funding from Pfizer to study smoking cessation. RW reports grants and personal fees from companies that develop and manufacture smoking cessation medications, outside the submitted work. MM has received unrestricted research funding from Pfizer and Rusan to study smoking cessation.

## REFERENCES

1. Dunbar MS, Davis JP, Rodriguez A, Tucker JS, Seelam R, D'Amico EJ. Disentangling Within- and Between-Person Effects of Shared Risk Factors on E-cigarette and Cigarette Use Trajectories From Late Adolescence to Young Adulthood. *Nicotine Tob Res.* 2019;21(10):1414-1422.

2. Kandel D. Stages in adolescent involvement in drug use. *Science*. 1975;190(4217):912-914.
3. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int. J. Epidemiol.* 2016;45(6):1866-1886.
4. National Academies of Science Engineering and Medicine. Public Health Consequences of E-Cigarettes (<http://nationalacademies.org/hmd/Reports/2018/public-health-consequences-of-e-cigarettes.aspx>). 2018.
5. Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S. The validity of self-reported smoking: a review and meta-analysis. *Am.J.Public Health*. 1994;84(7):1086-1093.
6. Chan GCK, Stjepanovic D, Lim C, et al. Gateway or common liability? A systematic review and meta-analysis of studies of adolescent e-cigarette use and future smoking initiation. *Addiction*. 2021;116(4):743-756.
7. Vanyukov MM, Tarter RE, Kirillova GP, et al. Common liability to addiction and "gateway hypothesis": theoretical, empirical and evolutionary perspective. *Drug Alcohol Depend*. 2012;123 Suppl 1:S3-17.
8. Shahab L, Beard E, Brown J. Association of initial e-cigarette and other tobacco product use with subsequent cigarette smoking in adolescents: a cross-sectional, matched control study. *Tob. Control*. 2021;30:212-220.
9. Munafo MR, Davey Smith G. Robust research needs many lines of evidence. *Nature*. 2018;553(7689):399-401.
10. Khouja JN, Suddell SF, Peters SE, Taylor AE, Munafo MR. Is e-cigarette use in non-smoking young adults associated with later smoking? A systematic review and meta-analysis. *Tob. Control*. 2021;30:8-15.
11. Bentivegna K, Atuegwu NC, Oncken C, DiFranza JR, Mortensen EM. Electronic Cigarettes Associated With Incident and Polysubstance Use Among Youth. *J. Adolesc. Health*. 2021;68(1):123-129.
12. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* 2003;32(1):1-22.
13. Khouja JN, Wootton RE, Taylor AE, Davey Smith G, Munafo M. Association of genetic liability to smoking initiation with e-cigarette use in young adults. *PLoS Med*. 2021;18(3):e1003555.
14. Cleveland HH, Wiebe RP. Understanding the association between adolescent marijuana use and later serious drug use: gateway effect or developmental trajectory? *Dev. Psychopathol.* 2008;20(2):615-632.
15. Beard E, West R, Michie S, Brown J. Association between electronic cigarette use and changes in quit attempts, success of quit attempts, use of smoking cessation pharmacotherapy, and use of stop smoking services in England: time series analysis of population trends. *BMJ*. 2016;354:i4645.
16. Meza R, Jimenez-Mendoza E, Levy DT. Trends in Tobacco Use Among Adolescents by Grade, Sex, and Race, 1991-2019. *JAMA Netw Open*. 2020;3(12):e2027465.
17. Walker N, Parag V, Wong SF, et al. Use of e-cigarettes and smoked tobacco in youth aged 14-15 years in New Zealand: findings from repeated cross-sectional studies (2014-19). *Lancet Public Health*. 2020;5(4):E204-E212.

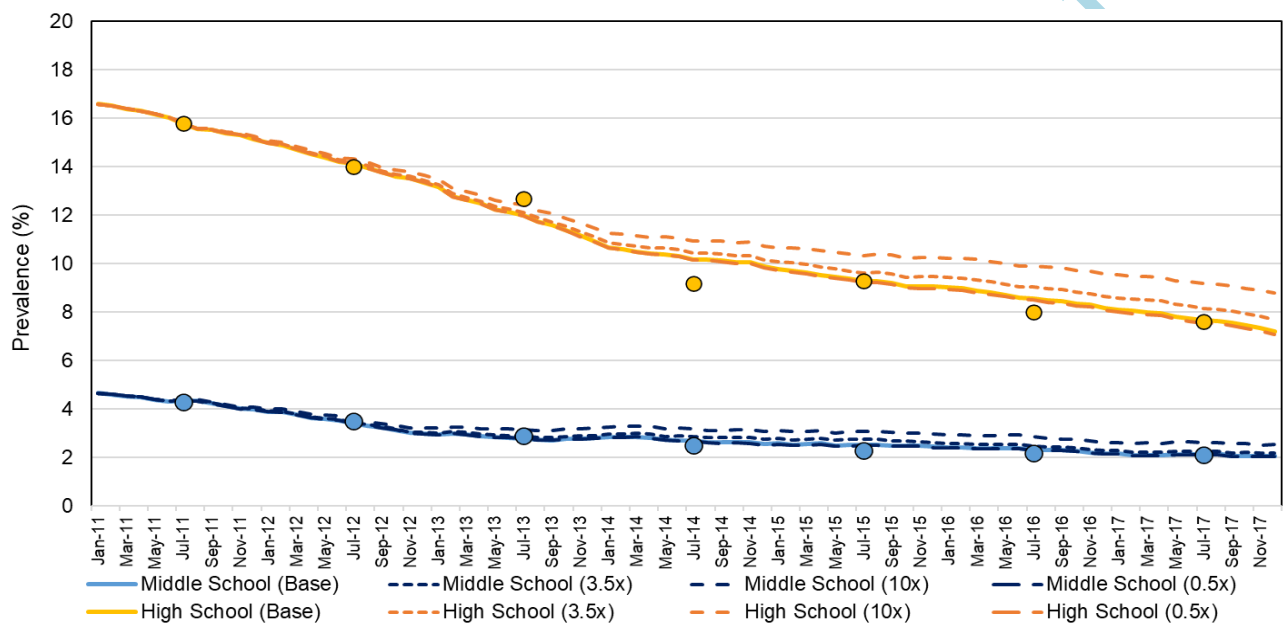


18. Rutter CM, Zaslavsky AM, Feuer EJ. Dynamic microsimulation models for health outcomes: a review. *Med. Decis. Making.* 2011;31(1):10-18.
19. Soneji S, Barrington-Trimis JL, Wills TA, et al. Association Between Initial Use of e-Cigarettes and Subsequent Cigarette Smoking Among Adolescents and Young Adults: A Systematic Review and Meta-analysis. *JAMA Pediatr.* 2017;171(8):788-797.
20. Beard E, Dienes Z, Muirhead C, West R. Using Bayes factors for testing hypotheses about intervention effectiveness in addictions research. *Addiction.* 2016;111(12):2230-2247.

Accepted Manuscript

**Figure 1: Observed\* and modelled past 30-day youth smoking prevalence in the USA 2011-2017, using the youth e-cigarette microsimulation model<sup>^</sup>**

\*Observed values (filled circles) come from the National Youth Tobacco survey 2011-2017; <sup>^</sup>The model, data and description can be found online (<https://osf.io/pycqi/>); briefly, the microsimulation consists of 50,000 agents, each of which represents an individual as defined by the characteristics relevant for the question (i.e., age, smoking status, vaping status), who at monthly intervals decide to take up smoking and/or vaping; the probabilities that govern these decisions are determined by the user (i.e. a multiplier that adjusts the probability of smoking uptake for agents that already vape and vice versa). Different postulated effect sizes (multipliers) for the strength of association of e-cigarette use with uptake of cigarettes is provided as odds ratios in brackets (x); the base model assumes no effect of e-cigarettes (solid line), broken lines indicate either a postulated positive (OR>1) or negative association (OR<1) between e-cigarette use and smoking uptake.



Accepted