STUDY PROTOCOL

Harmful health effects of e-cigarettes (vapes) and other non-tobacco oral nicotine products, trends in their use, and interventions: living protocol for a living evidence map and linked research digests

[version 1; peer review: awaiting peer review]

Gareth J Hollands[1], Ian Shemilt[1], Katy Sutcliffe[1], Claire Stansfield[1], Amanda J Sowden[2], James Thomas[1]

[1]EPPI Centre, UCL Social Research Institute, University College London, London, England, WC1H 0NS, UK
[2]Centre for Reviews and Dissemination, University of York, York, England, YO10 5DD, UK

v1 **First published:** 20 Jun 2025, **14**:609
https://doi.org/10.12688/f1000research.165927.1
Latest published: 20 Jun 2025, **14**:609
https://doi.org/10.12688/f1000research.165927.1

Abstract

There is a recognised need to further understand potential harms from the use of e-cigarettes and other non-tobacco oral nicotine products (ONPs), particularly among children and young people. This protocol describes a programme of work that aims to continually identify and classify the evolving international evidence base on the effects of these products on human health, trends in their use, and related interventions. In terms of its methods, we will produce and maintain a living evidence map, available in an online interactive format, and updated continually over the planned duration of the project (~5 years from registration). A wide range of electronic databases will be searched, with eligible articles including a focus on the use of e-cigarettes or ONPs in humans. Should satisfactory levels of machine performance be reached (by reference to human performance benchmarks), we intend to automate the bulk screening of articles for inclusion using Large Language Model (LLM) decisions. A similar approach will be applied to the automated coding of a range of characteristics of included articles, including their type of research, and the nature of harms and products being examined. In conjunction with the production of the living evidence map, we will produce regular descriptive surveillance reports (research digests) that will summarise the current contents of the map. Ongoing production and maintenance of this living evidence map and linked outputs will

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

enable potential users – including across research, policy, and charitable sectors – to more efficiently establish an overview of the landscape of relevant research and surveil key developments.

Keywords
vaping; vapes; e-cigarettes; oral nicotine products; health; harms; living evidence map; evidence synthesis

This article is included in the Addiction and Related Behaviors gateway.

Corresponding author: Gareth J Hollands (gareth.hollands@ucl.ac.uk)

Author roles: Hollands GJ: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Shemilt I**: Data Curation, Formal Analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – Review & Editing; **Sutcliffe K**: Conceptualization, Funding Acquisition, Investigation, Supervision, Writing – Review & Editing; **Stansfield C**: Data Curation, Investigation, Methodology, Software, Writing – Review & Editing; **Sowden AJ**: Conceptualization, Funding Acquisition, Investigation, Writing – Review & Editing; **Thomas J**: Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Resources, Software, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work has been commissioned by the National Institute for Health Research (NIHR) Policy Research Programme (PRP) for the Department of Health and Social Care (DHSC). It is funded through the NIHR Policy Research Programme contract with the EPPI Centre at UCL (Reviews facility to support national policy development and implementation, NIHR200701). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR or the DHSC. We thank DHSC colleagues for their critical feedback throughout the development of this protocol.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Hollands GJ, Shemilt I, Sutcliffe K *et al.* **Harmful health effects of e-cigarettes (vapes) and other non-tobacco oral nicotine products, trends in their use, and interventions: living protocol for a living evidence map and linked research digests [version 1; peer review: awaiting peer review]** F1000Research 2025, **14**:609 https://doi.org/10.12688/f1000research.165927.1

First published: 20 Jun 2025, **14**:609 https://doi.org/10.12688/f1000research.165927.1

Introduction

The development of this living evidence map is informed by the evolving research and policy landscape, both in the context of England and the United Kingdom (in which the work has been commissioned and will be conducted), and more widely. In this section we briefly summarise these considerations and the related scientific literature. We focus particularly on children, adolescents and young adults, as those groups are considered of highest priority in identifying emerging and escalating harms.

Trends in use of e-cigarettes (vapes) and other non-tobacco oral nicotine products

Recent years, especially since 2021, have seen a marked increase in England in the use of e-cigarettes, also termed vapes or electronic nicotine delivery systems (ENDS) [1].[1] Although the largest increases in persistent e-cigarette use have been seen in those with a history of regular smoking (reflecting their use as a smoking cessation tool), there has also been a substantial increase in adolescents and young adults, including those who have never smoked.[2] Prior to this relatively recent trend, e-cigarettes were mainly used by adults trying to quit smoking. While data from 2024 suggest that rates in 11-17 year olds appear to have now stabilised, levels of use remain significant, with 18% reporting having ever vaped and 7.2% reporting being current users.[1] This increase in prevalence has been linked to the wide range of single use (disposable) devices that have concurrently entered the market, as well as deliberately youth-friendly branding and advertising.[3] In recent survey data, more than half of those aged 11-17 who vape report doing so using a disposable device,[1] although following the announcement in January 2024 of an impending UK ban on disposable vapes, there has been a shift away from their use (from 63% to 35% amongst 16-to-24-year-olds).[4]

Alongside e-cigarettes, an increasing array of other non-combustible and non-tobacco oral nicotine products (ONPs) is becoming available, including pouches, lozenges, tablets, gummies and gum.[5,6] Many of these ONPs use synthetic nicotine (i.e., not derived from tobacco). Aside from variants licensed and intended for use in nicotine replacement therapy in smoking cessation, such as some types of lozenges and gums, consumer ONPs (as is the case for e-cigarettes) have not typically been approved for therapeutic purposes. These products have also been marketed towards use by children and young people, both in terms of branding and advertising, as well as the product forms aligning with confectionery products and featuring appealing flavours. The use of nicotine pouches is being monitored in the Smoking Toolkit study[7] which surveys usage of tobacco and nicotine products in those aged 16 and over in England, with data from October 2024 reporting low but increasing usage of nicotine pouches (current prevalence of 0.9%) relative to e-cigarettes (12.7%).

Evidence of harmful health effects of use

There is scientific consensus that the use of e-cigarettes and other ONPs is likely less harmful than is tobacco use, as well as strong evidence that they can be an effective smoking cessation tool for adult smokers.[8] However, there is also a growing literature documenting health harms. For example, a recent overview of reviews examining the potential harmful effects of e-cigarette use in people under 25 years old, found consistent associations between use and diagnosis or exacerbation of asthma, cough, depression, suicidal outcomes, and smoking, marijuana and alcohol use behaviours [Golder et al, unpublished report]. It is plausible that over time and with prolonged population use, incidence of known harms will escalate, and new, currently unknown harms will emerge. There are also many known potential harms for which there is currently a relative absence of evidence, including oral health, fertility, and brain development. Finally, the methodological quality of review-level evidence for harms of e-cigarette use has been assessed as of generally low quality suggesting that there is a need for continued development of the evidence base [Golder et al, unpublished report].

Current policy context in the United Kingdom

The work outlined in this protocol was commissioned broadly concurrently with significant developments in related policy in the United Kingdom, principally the need to gain further understanding of potential harms from the use of e-cigarettes, particularly among children and young people given the recent rise in youth vaping. The Tobacco and Vapes Bill 2024-25 was introduced to Parliament in 2024, with its stated aim being to create a smoke-free UK. It includes measures to regulate vapes and nicotine to reduce their appeal and accessibility to children. It will ban the advertising and sponsorship of all vapes and other nicotine products, as well as provide powers to regulate the flavours, packaging and shop displays of these products to prevent them from being deliberately promoted to children. The Bill will also enable licensing schemes and related enforcement for the retail of tobacco, vapes and other nicotine products. At the time of writing, the Bill passed its second reading in the House of Lords on Wednesday 23rd March 2025 and is approaching committee stage. Royal Assent is expected to be reached by the end of this Parliamentary session.

[1]E-cigarettes being the most commonly used term for these products in the scientific literature and our preferred term in this protocol.

Further UK governmental policy work on vaping beyond the Tobacco and Vapes Bill includes the single-use ban and the Vaping Products Duty. Under separate environmental legislation, the Department for Environment, Food and Rural Affairs (DEFRA) have laid regulations that will ban the sale and supply of any single-use vapes from 1 June 2025.[9] The ban will come into force on the same day across England, Wales, Scotland and Northern Ireland. Additionally, at the Autumn Budget 2024, the UK Government announced the Vaping Products Duty will come into force in October 2026, introducing a flat-rate excise duty of £2.20 per 10 ml to all vaping liquids to discourage young people from taking up vaping.

Scope of this work

Given its purpose to support policy development concerning actions to reduce or mitigate the potential harms of using e-cigarettes and other ONPs, this living evidence map will principally seek to describe the extent and nature of existing research that directly examines health impacts. However, its scope will also encompass key related areas, including evidence of trends in the use of e-cigarettes and ONPs including relationships with the use of tobacco products (e.g. in general use or as a smoking cessation aid); evidence of the public's perceptions of and discourse around the potential harmful effects of these products; and evidence on interventions targeting e-cigarette and ONP use. In parallel to other consumable products (most obviously tobacco products), a wide variety of interventions and policies may be applied to such products, and these can be categorised in many ways. A key commonly-drawn dimension is whether interventions are: i) targeted at or delivered directly to individuals or groups, or ii) are population-level interventions delivered at scale. The former includes pharmacological and behavioural support interventions.[10] The latter includes economic interventions such as taxes[11] and interventions that target physical environments at both micro and macro scales,[12] including those that alter the characteristics of products and related marketing and promotional activities.[13,14] Such interventions can be implemented within or outwith formal regulatory or legislative contexts.[15]

Overall purpose and aims

The primary purpose of this living evidence map is to support ongoing policy development processes in the United Kingdom that concern addressing potential harmful health impacts of the use of e-cigarettes and other non-tobacco ONPs, particularly in relation to children and young people. It will enable potential users - including stakeholders involved in commissioning this research, but also from wider government, research, policy, healthcare professional, and charitable sectors – to surveil the evolving international evidence base and more efficiently establish an understanding and overview of the landscape of relevant research, including its key characteristics. It will also underpin a linked wider programme of work including a suite of systematic reviews or other evidence syntheses. Given its intention to support policy development around harms of these products, the map principally seeks to describe the extent and nature of existing research that directly examines harmful health effects, but it will also encompass key cognate areas of the literature concerning trends in the use of these products, and interventions. Production of the map will be semi-automated (i.e. combining human and machine effort), with the intention to incrementally increase the degree of automation as far as is practicable. Potential benefits of increasing automation include enabling manual resources that are saved to instead be deployed towards additional aspects of map production or linked outputs, and substantial automation could increase the likelihood of the map being able to remain updated beyond the current timescale. It will be published as an open-access, online, interactive web database, and is initially intended to be produced and updated for a period of five years upon initial registration of this protocol.

The specific aims of the living map are as follows:

i) To continually identify and classify the evolving international evidence base of research studies on the effects of e-cigarette and other non-tobacco oral nicotine product use on human health;

ii) To directly inform the production of regular descriptive surveillance reports (termed 'research digests') over its duration that describe how the evidence base is evolving, including identification of emerging phenomena and of evidence gaps;

iii) To periodically support the efficient identification of evidence to be used in the conduct of systematic reviews or other evidence syntheses for prioritised research questions.

This protocol focuses mainly on the production of the living evidence map (i.e. Aim i), and to a lesser extent on the linked regular surveillance reports ('research digests') (i.e. Aim ii). In the section *'Additional evidence synthesis outputs'* we also outline how we anticipate the map will be used in relation to Aim iii. Because the nature of the intended outputs linked to Aim iii (systematic reviews or other evidence syntheses) is not yet specified and will be developed in consultation with stakeholders and reflecting evolving policy priorities, we will subsequently develop specific protocols to address this aspect.

Methods

We will produce a living evidence map to continually identify and classify the evolving body of published research evidence on the health impacts of e-cigarettes and other emerging non-tobacco oral nicotine products, and related bodies of evidence on trends in their use, and on interventions. Evidence maps are systematic and visual representations of the availability of evidence given a particular focus or domain relevant to a research question[16,17] and provide users with both a view of the broad research landscape and the possibility to focus on individual records of research studies [Shemilt et al, unpublished report]. They commonly consist of a framework of primary dimensions (rows and columns) and secondary dimensions as a set of filters, enabling exploration of the map's contents. Additionally, sectors of the map can be delineated to focus on specific topics or subsets of the overall map contents.

The 'living' aspect (of 'living evidence map') reflects our intention to keep this map continually updated with new research studies as they are published and to republish updated versions of the map at regular intervals during the project lifespan [Shemilt et al, unpublished report]. It also reflects the likelihood that key characteristics of the body of research included in the map may evolve over its lifetime, thus requiring changes in the scope of the map and how the evidence it contains is classified. Similarly, the purpose and aims of the map may evolve over time due, for example, to changes in the policy landscape it is designed to inform. Finally, the methods and procedures used to produce the map may also evolve based, for example, on embedded evaluations of automation tools that will be developed and deployed to assist its production. Correspondingly, this protocol is considered a living document and may be re-published in an updated separate version to incorporate any substantive changes to the protocol in the lifetime of the map. Any such potential changes will be discussed with commissioners prior to commencing production of each version of the living map.

The proposed methods for the living map, including our judgement that large-scale automation of screening and data extraction processes is likely feasible, have been informed by an initial exploratory developmental phase that involved production of, and collaborative discussion focused on, a rapid semi-automated evidence map [Hollands et al, unpublished report].

Eligibility criteria

The following criteria will determine whether reports of studies are included or excluded from the living evidence map:

Population: Humans, with an inclusive scope across children, adolescents and adults. It is noted that while there is currently particularly high policy interest in impacts on children, adolescents and younger adults (e.g. under 25 years of age), we do not at present intend to limit inclusion in the map to evidence that focuses on these groups (by for example, limiting the map to research that focuses solely, or to at least some extent, on children and young people). We will exclude reports of studies in animals and in vitro or other types of studies that lack evaluation or assessment in humans.

Exposure, intervention, or topic focus: Use of e-cigarettes (i.e. nicotine and non-nicotine vapes or vaporisers or electronic nicotine delivery systems (ENDS)) or other non-tobacco oral nicotine products (ONPs) (e.g. nicotine pouches, lozenges, tablets, gummies, gum, spray). For reports of intervention studies, we will include any intervention or policy that explicitly aims to alter the use of these products. These may aim to reduce the use of, or the harms linked directly to, e-cigarettes and ONPs, or aim to increase use of these products whether for therapeutic (e. g. smoking cessation) or other reasons, such as commercial promotional activities or alterations to product design to increase appeal and use.

Outcomes: Any quantitative or qualitative measure or assessment of e-cigarette or ONP use, or of any health effects or harms that are directly or causally linked to e-cigarette or other ONP use (either as specified or as is considered plausible). This encompasses those relating to physical health (such as concerning respiratory, cardiovascular, oral, and neurological/developmental harms, as well as injuries and accidents), mental health (including depression, anxiety, and suicide), and behavioural impacts on substance use. The latter includes effects on concurrent or subsequent use of other nicotine and tobacco products (e.g. cigarette smoking, or the use of oral and heated tobacco products), as well as of alcohol and other recreational drugs including the vaping of cannabis. We will also include quantitative or qualitative measures of perceptions of the harms of e-cigarette or ONP use.

Study designs: We will include all empirical research and will not limit inclusion by study design. We will include both primary and secondary research (e.g. systematic reviews) designs. Given an intended function of the map is to aid in identifying emerging phenomena related to harms, we will include descriptive designs considered to be less methodologically robust, such as case reports and case studies, but which may provide useful early evidence. We will exclude non-empirical research papers including commentaries, editorials, and non-systematic reviews.

Publication types: Peer-reviewed journal articles published since 2003 (the year in which e-cigarettes became commercially available, preceding their estimated entry into Europe in 2005[18]) will be included, including those

reporting protocols of otherwise-eligible empirical research. We will exclude preprints, commentaries, editorials, correspondence, conference abstracts, theses and dissertations. We will limit to English language reports only due to both the substantial numbers of records to be processed (any of which may require additional manual reading and oversight) and our intent to extensively use automated processes that may perform differently dependent on the language of the text.

Identifying the evidence

Initially, the following electronic databases will be searched, containing scholarly research across health care, behavioural and social sciences and multi-disciplinary fields. The following list may be modified over the course of the map's development subject to evaluation of how effectively these sources contribute to identification of eligible reports:

- Applied Social Sciences Index and Abstracts (ASSIA) (ProQuest)
- CINAHL (EBSCO)
- Cochrane Library - Databases of Systematic Reviews and controlled trials
- Database of promoting health effectiveness reviews (DoPHER)
- Embase (OVID)
- Emerging Sources Citation Index (Web of Science)
- Epistemonikos
- Health Management Information Consortium (OVID)
- MEDLINE (OVID)
- OpenAlex
- PsycInfo (OVID)
- Science Citation Index (Web of Science)
- Social Science Citation Index (Web of Science)
- The Trials Register of Promoting Health Interventions (TRoPHI)

The search strategies for the full set of databases will be developed in collaboration with an Information Specialist. The MEDLINE search terms have been informed by those published for other relevant reviews,[10,19,20] [Golder et al, unpublished report]. This initial search will involve conventional Boolean searching using terms for e-cigarettes and non-tobacco oral nicotine products. It will exclude animal-only studies and records classed as preprints, commentaries, editorials, correspondence, conference abstracts, theses and dissertations. We will use a range of topic relevant terms and synonyms and search the title and abstract fields of records as well as controlled vocabulary within individual databases, such as Medical Subject Headings (MeSH).

OpenAlex is an open-access knowledge graph comprising ~250 million bibliographic records of research articles (reports) from across science, connected in a large network graph of conceptual, citation and other (e.g. author) relationships. OpenAlex is automatically and continuously updated with new records as new research reports are published online. We have primarily developed OpenAlex tools in EPPI Reviewer[21] to perform regular automated searches of the OpenAlex knowledge graph to support continual, semi-automated updating of living evidence maps and living systematic reviews. OpenAlex tools enable three different kinds of searches of the OpenAlex knowledge graph: custom search, network graph search and auto-update search. Custom search features will first be used to run the Boolean search of OpenAlex as part of the initial search described above, and we will investigate options of using focused or sensitive Boolean searches of this source.

Subsequently, our continual updating searches will comprise all three kinds of OpenAlex searches: custom search (updated), network graph search, and auto-update search. A network graph search will retrieve all records that are connected, in OpenAlex, to a specified set of 'seed' records, via a specified set of network graph relationships, on the date of the search. Specified 'seed' records will be the accumulating corpus of unique eligible reports already selected for inclusion in the map (which have either been retrieved from OpenAlex or matched to a corresponding OpenAlex record from another source database record). Specified network graph relationships will be: (a) one-step forwards (i.e. records that cite 'seed' records) or backwards (records that are cited by 'seed' records) citation relationships (i.e. equivalent to citation searching); and (b) a one-step forwards ('recommended by' seed records) 'related publications' relationship [2]. An auto-update search will likewise be 'seeded' by the accumulating set of unique eligible reports already selected for inclusion in the map. These 'seed' records will be subscribed to our novel machine learning 'recommender' model (the 'auto-update model'),[22] which will automatically score all records (reports) newly added to the OpenAlex dataset each ~1 month in its 'snapshots' and recommend those most likely to be eligible for inclusion in this map. We will select the top 250 ranked records by auto-update model score from each 'snapshot'. Initially, we will run updated custom and network graph searches, and import records from all three kinds of OpenAlex searches into EPPI Reviewer, every four months, i.e. at the beginning of each updating cycle, following the publication of each updated version of the map.

Selection procedure

Each time the initial database or continual updating searches are run, all bibliographic records will be imported into EPPI Reviewer. Any records of articles published before 1st January 2003 will be removed. Duplicates (both within the set of searches, and in relation to records already existing in the map) will be semi-automatically identified and discarded using 'manage duplicates' tools. The process for then selecting eligible records for inclusion in the map is as follows.

First, in a pilot phase, in order to inform development of operationalisations of our eligibility criteria that are able to be applied consistently, a random sample of 100 title-abstract records (and corresponding full-text reports when needed) drawn from the full search results will be screened for inclusion/exclusion by two human reviewers independently and their decisions compared and discussed to reach final decisions. Additional samples of 100 records (and corresponding full-text reports) will be screened by both reviewers until a high degree of agreement on inclusion/exclusion decisions (90% or more) is achieved. At this stage, we will also discuss any apparent 'borderline' eligibility examples. The objectives of the pilot phase will be to identify revisions to screening codes, as well as to foster consistency between researchers, thus ensuring we apply criteria and screening processes that minimise the exclusion of potentially important records and minimise inclusion of clearly irrelevant records.

Second, manual screening will be undertaken by one trained researcher ('single screening') to provide a set of records for use when developing and evaluating a subsequent automated screening workflow. For this, a set of title-abstract (plus corresponding full-text article) records – comprising a minimum of 1000 records - will be drawn from the full search results at random and screened to assess eligibility, with the further option of referring records for a 'second opinion' if eligibility is judged uncertain. 'Second opinions' will be resolved by team discussion and consensus, involving two or more researchers (including the original screener). We will aim to retrieve and upload corresponding full-text articles (usually PDFs) for each of these records (reports) using a manual search via the DOI or OpenAlex link in each record. The set of human-screened title-abstract (and corresponding full-text article) records will then be considered the (quasi) 'gold-standard' benchmark for assessing performance of subsequent automated inclusion/exclusion screening, and will enable calculation of metrics such as % agreement and inter-rater reliability metrics that adjust for chance in agreements of inclusion and exclusion decisions (for example, Cohen's kappa). While we are not necessarily aiming for such a stringent overall kappa value when assessing performance, the benchmark sample size of $\geq$1000 is justified based on the formula for the standard error of kappa with the aim to achieve a kappa of at least 0.9 (typically interpreted as almost perfect agreement), with a conservative assumed base rate of 75% of records being included [3], and with a confidence interval no wider than 0.05 to exclude the possibility of any less than strong agreement.[23] This benchmark sample size also exceeds sample sizes typically used in inter-rater agreement studies.[24]

Third, for developing and evaluating an automated screening workflow, in accordance with a standard procedure,[25] the overall set of manually screened records and their corresponding screening meta-data will be split into three samples of different sizes and compositions, using stratified random assignment. We will develop a series of Large Language Model

[2]OpenAlex 'related publications' (records) are recommended by 'seed' publications (records) based on their ranking (top ranked are those which score highest) on a composite metric that combines all of the various network graph relationships available within the OpenAlex knowledge graph.

[3]An initial sample of 100 records screened in early-stage development of this work suggested 62% of records would be included and so using this estimate of 75% is more conservative in requiring a larger sample size.

(LLM) prompts to elicit machine inclusion/exclusion screening decisions via ChatGPT 4o [4][26] functionality embedded within EPPI Reviewer software.[21] The first sample will be used in early rounds of developing these LLM prompts, aimed at identifying a set of prompts (with or without a decision algorithm) that maximises recall as much as possible (with a target of 0.95 recall). The second sample will be used to further develop and finalise the set of prompts (with or without a decision algorithm), aimed at maximising precision without compromising recall in relation to the approximate level attained in the first sample. The third sample will be used to evaluate the finalised set of prompts (with or without a decision algorithm).

Should we achieve satisfactory performance for automated inclusion/exclusion screening, we intend to subsequently automate the bulk screening of all records that are identified by searches, to determine their inclusion into or exclusion from the map. If we are unable to automate screening satisfactorily then we will amend the protocol accordingly to employ an alternative workflow for screening that applies a greater degree of manual screening and oversight (i.e. a semi-automated workflow). Furthermore, should there be any substantive changes to our planned processes in the lifetime of the map, that suggest that we can no longer consider our corroboration of this screening process to be valid, a comparable testing process will be newly conducted. Such possible substantive changes could involve, for example, alterations of inclusion/exclusion criteria or the nature of the search strategies, or advances or other changes in the computing technologies we deploy. As previously mentioned, reflecting that this document is a living protocol, we will also consider whether any such changes made to our processes merit publishing an updated version of this protocol.

In addition, once we have accumulated a sufficient number of manual screening decisions, we will investigate training, calibrating, evaluating and (contingent on evaluated performance) deploying a binary machine learning classifier, designed to distinguish between records (reports) of eligible (included) and ineligible (excluded) studies, using machine learning tools in EPPI Reviewer. This classifier would be calibrated and deployed to automatically exclude (discard) further retrieved records (reports) prior to screening, reducing the numbers of records to be screened with minimal risk of losing relevant potentially eligible reports.

Classifying the evidence

We will aim to classify included reports in relation to the following eight characteristics, each represented by sets of coding options, and hereforth termed ‘code sets’. The code sets below are provisional and illustrative, being informed by initial discussions with commissioners, rapid exploratory mapping work to feed into those discussions, and examination of relevant literature such as existing systematic reviews of evidence on e-cigarette harms. The formulation and categorisation of the characteristics and code sets that we focus on may evolve over the lifetime of the map. Developments to code sets may stem from the map’s contents (e.g. new categorical distinctions emerging as new evidence is encountered, or improved operationalisations arising from resolving queries and inconsistencies) or from top-down or external influences (e.g. applying new conceptual knowledge as it is encountered via wider discussion, or changes in research and policy landscapes and priorities).

Code set 1 - Broad focus or topic

- Direct effects or harms of e-cigarette or ONP use (excluding smoking/tobacco product use)
- Longitudinal trends in e-cigarette or ONP use and/or its relationship with concurrent or subsequent smoking/tobacco product use
- Public discourse or perceptions concerning effects or harms of e-cigarette or ONP use
- Effects of interventions or policies to reduce or increase e-cigarette or ONP use

Code set 2 - Effects or harms of e-cigarette or ONP use *(where relevant - coding applied only to those coded in 1(Broad focus or topic) as category ‘Direct effects or harms of e-cigarette or ONP use’)*

- Respiratory (e.g. lung injury, asthma etc)
- Cardiovascular

[4]Or a newer model, should one become available that is better suited to this task.

- Cancer
- Immune/allergy
- Injuries
- Oral health
- Fertility
- Pregnancy and perinatal
- Infectious disease
- Mental health
- Biological/physiological/cellular
- Substance use (excluding smoking/tobacco product use)
- Long-term population-level impacts (e.g. mortality)
- Other (specify)

Code set 3 - Broad class of research

- Primary research report
- Study protocol
- Systematic review or other evidence synthesis

Code set 4 - Product focus

- E-cigarettes (nicotine and non-nicotine)
- Nicotine pouches
- Gummies
- Gum
- Lozenges
- Tablets
- Spray

Code set 5 - Population

- Exclusively children and/or adolescents (<18 years)
- Population includes, or is exclusively, adults (18+ years)

Code set 6 - Country

- UK
- Non-UK Europe
- USA
- Canada
- Mexico
- Australia
- New Zealand
- China
- Japan
- Other (specify)

Code set 7 - Detailed type of intervention or policy to reduce or increase e-cigarette or ONP use *(where relevant - coding applied only to those coded in 1(Broad focus or topic) as category 'Effects of interventions or policies to reduce or increase e-cigarette or ONP use')*

Individually/group delivered to reduce use

- Pharmacological
- Behavioural support

Population-level delivered to reduce use

- Economic/fiscal (with or without legislation or regulation)
- Physical environmental including characteristics of products and (non-economic) promotions (with or without legislation or regulation)

Individually/group delivered to increase use (e.g. for therapeutic or marketing purposes)

- Pharmacological
- Behavioural support

Population-level delivered to increase use (e.g. for therapeutic or marketing purposes)

- Economic/fiscal (with or without legislation or regulation)
- Physical environmental including characteristics of products and (non-economic) promotions (with or without legislation or regulation)

Code set 8 - Detailed type of study design (supplementing code set 3) [5]

Primary research

- Cross-sectional survey
- Other descriptive (e.g. qualitative, case study or case report)
- RCT (i.e. control/comparison to a group(s) who do not receive intervention for vaping cessation), including cluster-randomised
- Longitudinal design (e.g. simple pre/post comparison, interrupted time series, cohort or repeat cross-sectional)
- Other quasi-experimental design (e.g. non-randomised controlled trial, controlled before-and-after)

Secondary research

- Systematic review
- Overview of reviews
- Evidence map

We intend to automate the classification (coding) of included reports, using a similar approach to that applied for automating screening to determine inclusion into the map. For each code set, in an initial pilot phase, we will develop mutually exclusive codes that are considered both appropriately granular to be useful and sufficiently comprehensive to apply to the range of research we encounter. When developing code sets, records will initially be coded in increments of ~5-10 articles by two reviewers working independently. Further increments of articles will be extracted by both reviewers until no substantive changes to the coding structure are required to accommodate the evidence, and agreement is maximised as far as possible. For developing and evaluating the automated coding workflow, a similar corroboration process will be used for as is described for screening (in 'Selection procedure'), albeit less extensive, in part because we consider the accuracy of coding to be of a lower relative priority than determining the accuracy of screening. This is because errors in screening can result in records being lost permanently from the map's confines, while errors in coding will not prevent the record from remaining findable in the overall map (such as when the map's contents are searched or explored in different ways) and may be relatively minor for any given record (e.g. an error in relation to one code set only). Furthermore, coding of each record's many characteristics will likely require more human effort per record than will screening, meaning generating benchmark samples of manually-coded records will be more resource-intensive. The combination of these factors is reflected in each code set that we develop and apply being evaluated in relation to a smaller manually-coded benchmark sample, the size of which will be contingent on the time available and observed inclusion rates, but including a minimum of 100 records which, if possible, will include at least some examples of each code within the given code set. The levels of agreement for the automated coding workflow – which may vary substantially between the code sets - will be reported to allow the map's users to judge the likely reliability of its classifications. As for screening processes (see 'Selection procedure'), should we make changes to our automated coding processes that mean that we can no longer have sufficient confidence in our prior testing of them, we will conduct further tests for corroboration purposes.

We will initially prioritise effort towards developing and implementing code sets 1-3 which are considered to be of immediate priority as well as likely to be highly feasible to code given current tools. We intend to incrementally implement further code sets using a similar process and corroborate the performance of automating each. We will also consider supplementing the outlined standard processes of identifying evidence by populating the map with additional key exemplar papers as identified by the research team, collaborators, and advisory team, and related academic and policy networks, particularly in the early stages of its development before the existing literature is comprehensively represented via the standard processes. Once an initial version of the map (v1) is complete and functioning as expected, we will seek and address feedback from stakeholders – including in relation to the map's content as well as its presentation or useability

[5]We do not intend to systematically extract risk of bias or quality appraisal information for included records. It is possible, however, that we may develop an approach to annotating notably 'high-quality' studies (or bodies of studies) that are included in the map, such as through using known study design characteristics to infer likely explanatory power (e.g. systematic reviews identifying evidence of sufficient certainty, large-scale observational studies, randomised controlled trials).

- and proceed to develop further updated versions on a regular basis (3 times a year i.e. approximately every 4 months [6]).

Implementing the protocol as currently described depends on whether we are able to substantially automate the map, and on our current understanding of its likely scope. However, if, for example, a lesser degree of automation proves possible than is anticipated (in which case we will consider whether additional manual coding is possible and justified, including relative to its likely importance to users), or ongoing input from stakeholders (including the wider research team, the Scientific Advisory Panel (SAP), and the commissioners of this work) prompts changes in the map's scope or content, then the outlined procedures may necessarily change to reflect changing priorities and resources. Any substantial deviations from the current version of the protocol will be recorded in updated versions of the protocol.

Production of linked research digests

In conjunction with the production of the living evidence map, we will produce regular descriptive surveillance reports (research digests) three times a year i.e. approximately every 4 months, reflecting the planned frequency of updates of the map itself. These research digests will consist of a standardised basic report to summarise the current contents of the living evidence map. While the exact content and format of these reports will develop iteratively in consultation with commissioners of this review and other stakeholders, we anticipate it will principally tabulate frequencies of data and correspondingly describe the nature of the evidence as follows.

In a series of tables, we will tabulate and cross-tabulate frequencies of reports included in the map, and the data extraction codes that have been applied to them (i.e. with each code set providing a categorical variable(s)). Coding data will be systematically presented for all code sets individually (see 'Classifying the evidence' for these proposed code sets), supplemented by selected cross-tabulations of code sets deemed to be of highest priority or interest.

Taking code sets 1-5 as illustrative examples, this could involve reporting, in a series of tables, the numbers of reports that focused broadly on direct effects or harms (relative to other broad types of reports), which harms those reports focused on, how many were primary research, how many focused on e-cigarettes, and how many focused on children and adolescents. Cross-tabulations of these data in additional tables would therefore allow, for example, identification of how many primary research reports focused on respiratory harms in children and adolescents linked to e-cigarette use. Tables will be accompanied by narrative explanation of the numbers, including highlighting notably large or emerging clusters of reports with particular combinations of characteristics, or their absence (i.e. apparent evidence gaps). Additional complementary information could be derived from direct comparison to the comparable data in previous reports, thus enabling mapping of gradual trends in the development of the evidence base, or the initial appearance of new clusters of evidence.

Additional evidence synthesis outputs

The continual production of the living evidence map (and corresponding research digests) will also underpin and inform a linked wider programme of less frequent but more detailed evidence synthesis work. Because the nature of the intended outputs linked to our Aim iii (systematic reviews or other evidence syntheses) is not yet specified and will be developed in consultation with stakeholders and reflecting evolving policy priorities, we will subsequently develop specific protocols to address these aspects. However, here we provide a brief indication of how we anticipate this aspect could develop.

At approximately 12-month intervals, we will produce a report of a systematic review or other evidence synthesis that will likely involve drawing substantially (but not necessarily exclusively) on evidence from the map. The research questions to be examined will be informed by current policy priorities as well as the contents of the map, and collectively agreed with the commissioners of the review. As illustration, the types of more detailed evidence synthesis outputs presented in these reports could include:

- Formal systematic or rapid reviews if, for example, there was a) a clear and important policy question, and the map's contents suggested the existence of b) a coherent body of relevant studies, and c) an evidence base of manageable scope and size. The living evidence map could both help to corroborate b) and c), as well as support rapid and efficient identification of at least some of the relevant evidence in conducting an agreed review

[6]In terms of timing, we intend to produce the first version of the map (v1) within an approximately six-month period subsequent to approval of this protocol by its commissioners and its public registration. This is because it is likely the initial development and corroboration of processes will take additional time. However, following that, we intend that each subsequent full version of the map would be published in accordance with a regular four-month cycle i.e. v2 would be published approximately four months after v1, and so forth.

- Higher-level or overarching evidence synthesis products such as interactive evidence portal(s) directly stemming from the contents of the map and with some additional assessment of evidence strength, summarising agreed key aspects of the map's contents but without requiring end users to engage with the underlying studies[27]

- More detailed additional coding, assessment and/or synthesis of evidence within specific priority aspects or areas of the living map itself

In supporting the development of these kinds of more detailed evidence synthesis outputs, the living evidence map could help identify viable research questions from observing noteworthy clusters of evidence. Furthermore, should a research question be developed independently of the map, the map could then allow us to ascertain whether any proposed evidence synthesis activities are likely able to draw on sufficient evidence to make them feasible and useful. In principle the existence of an up-to-date map should therefore reduce the likelihood of spending undue time reviewing in detail parts of the map (or aspects of research questions) that have already been well addressed by previous research, or of conducting uninformative or inefficiently-focused reviews of few studies. A final key benefit of a living evidence map is that it should increase efficiency in conducting any subsequent evidence syntheses. While it will not guarantee absolute comprehensiveness i.e. containing all eligible records pertaining to a given evidence synthesis or research question, it is likely that at least a substantial proportion of the relevant evidence will already be contained within the map and so will also have been annotated with information about its characteristics. Such bodies of known and annotated evidence will therefore provide a strong foundation to support subsequent evidence syntheses.

Stakeholder engagement

The work described in this protocol will be developed in an ongoing collaboration with commissioners in The Department of Health and Social Care (DHSC). The nature of this collaboration will develop iteratively but we anticipate will involve regular meetings to discuss progress and resolve issues that arise with the living map as well as the development of future linked outputs. Feedback will also be sought on whether any changes to the scope and policy relevance of the map or linked digests (and relatedly to the protocol) may be necessary. Regular meetings will typically coincide with the timing of the production cycle of living map updates and linked digests (i.e. approximately every four months).

Additionally, we will convene a Scientific Advisory Panel (SAP) with a membership external to both the research team and to DHSC. Its membership is yet to be confirmed and will be subject to agreement with DHSC, but the expectation is that it will comprise at least two academic researchers active in e-cigarette and/or relevant smoking-related research, and at least one member each from: i) a relevant charity or other third-sector organisation, and ii) the wider public (i.e. not employed in a role in this area), with all members ideally having a high level of knowledge or direct experience of the product landscape, related trends, or the relevant scientific literature particularly in the context of England and/or the wider UK. The purpose of the SAP will be to generally advise and respond to the research team's requests for feedback in their capacity as topic experts in relation to the focus of the map. This will include identifying whether there are new products or trends the map will need to pick up (as this is a rapidly evolving area), discussing specific technical queries arising from the living evidence map or linked outputs (e.g. whether specific boundary examples of reports should be deemed to meet our inclusion criteria, or how specific reports should most appropriately be categorised within our coding structure), and providing feedback on the general useability (e.g. accessibility, presentation and functioning) of the map itself. Such discussions will also likely involve discussion of the structure of the map and its coding schemes and whether they could be further optimised by, for example, reconfiguring categories, and making clarifications or changes to the wording of terms or definitions. More specific terms of reference (e.g. if regular SAP group meetings will be held and if so, how often, and/or if feedback will be sought largely on an ad hoc individual basis only) will be developed in discussions with DHSC and the members of the SAP.

Dissemination activities

We envisage the map may be useful to multiple potential audiences and users including stakeholders directly involved in commissioning and developing the planned research, as well as wider research, policy, practitioner, and charitable sector communities. Therefore, beyond the directly commissioned outputs, we intend to disseminate the existence of this work through academic and other stakeholder networks including presentations and social media posts. We will also consider producing academic articles (e.g. describing the methodological development processes or aspects of the map's content).

Equity considerations

Although this project does not include a specific focus on equity issues, we have planned to extract and report equity-related data on population characteristics as follows. First, in the living evidence map, we plan to classify (code) included studies in terms of the degree to which their populations involve children and/or adolescents (<18 years). Second, in our

linked research digests, we will report the latter data as part of our plan to present coding data for all code sets. This approach to considering equity issues in this project was informed by, and aligns with, developing EPPI Centre guidance on this topic.[28]

Ethical issues relating to this work

This is secondary research that only uses existing articles and with no new data being collected.

Study registration

The protocol for the living evidence map will be registered on the Open Science Framework at https://osf.io/8qbwv/.

Software availability

Not applicable for this article.

Data availability

No data are associated with this article.

References

1. Action on Smoking and Health: **Use of vapes (e-cigarettes) among young people in Great Britain (ASH Fact Sheet; July 2024).** 2024.
Reference Source
2. Jackson SE, Tattan-Birch H, Shahab L, *et al.*: **Trends in long term vaping among adults in England, 2013-23: population based study.** *BMJ.* 2024; **386**: e079016.
Publisher Full Text
3. Office for Health Improvement & Disparities: **Youth vaping: call for evidence analysis.** 2023.
Reference Source
4. Jackson SE, Shahab L, Tattan-Birch H, *et al.*: **Changes in vaping trends since the announcement of an impending ban on disposable vapes: A population study in Great Britain.** *Addiction.* 2025.
Publisher Full Text
5. Cornacchione Ross J, Kowitt SD, Rubenstein D, *et al.*: **Prevalence and correlates of flavored novel oral nicotine product use among a national sample of youth.** *Addict. Behav.* 2024; **152**: 107982.
Publisher Full Text
6. Dennison Himmelfarb CR, Benowitz NL, Blank MD, *et al.*: **Impact of Smokeless Oral Nicotine Products on Cardiovascular Disease: Implications for Policy, Prevention, and Treatment: A Policy Statement From the American Heart Association.** *Circulation.* 2025; **151**(1): e1–e21.
Publisher Full Text
7. Buss V, Kock L, West R, *et al.*: **Trends in electronic cigarette use in England (16th January 2025 update).** 2025.
Reference Source
8. Lindson N, Butler AR, McRobbie H, *et al.*: **Electronic cigarettes for smoking cessation.** *Cochrane Database Syst. Rev.* 2025; **1**: CD010216.
Publisher Full Text
9. Department for Environment, Food & Rural Affairs: **Single-use vapes ban: what businesses need to do.** 2025.
Reference Source
10. Butler AR, Lindson N, Livingstone-Banks J, *et al.*: **Interventions for quitting vaping.** *Cochrane Database Syst. Rev.* 2025; **1**: CD016058.
Publisher Full Text
11. Chaloupka FJ, Yurekli A, Fong GT: **Tobacco taxes as a tobacco control strategy.** *Tob. Control.* 2012; **21**(2): 172.
Publisher Full Text
12. Swinburn B, Egger G, Raza F: **Dissecting Obesogenic Environments: The Development and Application of a Framework for Identifying and Prioritizing Environmental Interventions for Obesity.** *Prev. Med.* 1999; **29**(6): 563–570.
Publisher Full Text
13. Hollands GJ, Bignardi G, Johnston M, *et al.*: **The TIPPME intervention typology for changing environments to change behaviour.** *Nat. Hum. Behav.* 2017; **1**: 0140.
Publisher Full Text
14. Marteau TM, Hollands GJ, Nambiar D, *et al.*: **Changing behaviour at scale to prevent NCDs.** Banatvala N, Bovet P, editors. *Noncommunicable Diseases: a Compendium.* Abingdon (UK): Routledge; 2023; pp. 347–354.
15. Marteau TM, Mantzari E, Hollands GJ: **Do nudges need a regulatory push? Comparing the effectiveness and implementation of exemplar nudge (size-based) and non-nudge (price-based) dietary interventions.** *Soc. Sci. Med.* 2025; **373**: 118004.
Publisher Full Text
16. Campbell F, Tricco AC, Munn Z, *et al.*: **Mapping reviews, scoping reviews, and evidence and gap maps (EGMs): the same but different— the "Big Picture" review family.** *Syst. Rev.* 2023; **12**(1): 45.
Publisher Full Text
17. Shemilt I, Hollands GJ, Stansfield C, *et al.*: *Debunk, Inform, Avoid? Debunking vaccine-related misinformation: a rapid evidence review. Report produced for the International Public Policy Observatory (IPPO).* London: EPPI Centre, UCL Social Research Institute, University College London; 2022.
Reference Source
18. Public Health England: **Electronic cigarettes.** 2014.
Reference Source
19. Hartmann-Boyce J, Tattan-Birch H, Brown J, *et al.*: **Oral nicotine pouches for cessation or reduction of use of other tobacco or nicotine products.** *Cochrane Database Syst. Rev.* 2025; **2**: CD016220.
Publisher Full Text
20. Hollands GJ, Naughton F, Farley A, *et al.*: **Interventions to increase adherence to medications for tobacco dependence.** *Cochrane Database Syst. Rev.* 2019; **8**: CD009164.
Publisher Full Text
21. Thomas J, Graziosi S, Brunton J, *et al.*: *EPPI-Reviewer: advanced software for systematic reviews, maps and evidence synthesis [Computer software].* London: EPPI Centre; 2023.
22. Tenti P, Thomas J, Peñaloza R, *et al.*: **ContReviews: A content-based recommendation system for updating Living Evidences in health care.** *Knowl.-Based Syst.* 2025; **311**: 112981.
Publisher Full Text
23. Wallrich L: **Assessing interrater agreement in meta-analyses – What sample size do I need?** 2023.
Reference Source
24. Han O, Tan HW, Julious S, *et al.*: **A descriptive study of samples sizes used in agreement studies published in the PubMed repository.** *BMC Med. Res. Methodol.* 2022; **22**(1): 242.
Publisher Full Text
25. Shemilt I, Thomas J, Churchill R, *et al.*: **Generative artificial intelligence tools versus conventional screening by humans for selecting eligible study reports for evidence synthesis: A living study within a review (SWAR) [Project registration and materials].** 2025.
Reference Source
26. OpenAI: **ChatGPT-4o [Large language model].** 2024.
27. White H: *The Strategic Use of Evidence and Gap Maps to Build the Evidence Architecture. CEDIL Methods Working Paper 5.* London and Oxford: Centre of Excellence for Development Impact and Learning (CEDIL); 2021.
Reference Source

28. Shemilt I, Hollands G, D'Souza P, *et al.*: **Considering and reporting methods for investigating health equity issues in systematic reviews and other evidence syntheses. (Version 1.0. October 2024).** London: EPPI Centre, Social Science Research Unit, UCL Institute of Education, University College London; 2024.
Reference Source