

Viewpoint

# Shopping Data for Population Health Surveillance: Opportunities, Challenges, and Future Directions

Alisha Suhag, BA, BSc, MSc, PhD; Romana Burgess, MMath, PhD; Anya Skatova, BA, MSc, PhD

Digital Footprints Lab & Medical Research Council Integrative Epidemiology Unit, MRC Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom

**Corresponding Author:**

Alisha Suhag, BA, BSc, MSc, PhD  
Digital Footprints Lab & Medical Research Council Integrative Epidemiology Unit  
MRC Integrative Epidemiology Unit, University of Bristol  
Oakfield House United Kingdom  
Bristol BS8 2BN  
United Kingdom  
Phone: 44 (0)117 331 0098  
Email: [alisha.suhag@bristol.ac.uk](mailto:alisha.suhag@bristol.ac.uk)

## Abstract

The growing ubiquity of digital footprint data presents new opportunities for behavioral epidemiology and public health research. Among these, supermarket loyalty card data—passively collected records of consumer purchases—offer objective, high-frequency insights into health-related behaviors at both individual and population levels. This paper explores the potential of loyalty card data to strengthen public health surveillance across 4 key behavioral risk domains: diet, alcohol, tobacco, and over-the-counter medication use. Drawing on recent empirical studies, we outline how these data can complement traditional epidemiological data sources by improving exposure assessment, enabling real-time trend monitoring, and supporting intervention evaluation. We also discuss critical methodological challenges, including issues of representativeness, data integration, and privacy, as well as the need for robust validation strategies. By synthesizing the current evidence base and offering practical recommendations for researchers, this paper highlights how loyalty card data can be responsibly leveraged to advance behavioral risk monitoring and support the adaptation of epidemiological practice to contemporary digital data environments.

*J Med Internet Res* 2025;27:e75720; doi: [10.2196/75720](https://doi.org/10.2196/75720)

**Keywords:** epidemiology; shopping data; health behaviors; diet; smoking; alcohol; surveillance; digital footprint data

## Introduction

Noncommunicable diseases (NCDs) account for over two-thirds of global mortality [1]. Addressing 4 primary modifiable risk factors—tobacco use, unhealthy diet, physical inactivity, and alcohol consumption—could prevent up to 80% of major NCDs [2]. However, monitoring these evolving risk factors at scale remains a major challenge in epidemiology.

Traditional data sources—such as self-reported surveys, health registries, and administrative records—are crucial for population health monitoring but increasingly limited by high costs, slow survey cycles, and reporting biases (eg, recall and social desirability), reducing their reliability for tracking dynamic health behaviors. Participation in cohort studies and epidemiological surveys is also declining, raising concerns

about selection bias and representativeness [3,4]. This is especially critical in research on lifestyle-related risk factors, where nonresponse often correlates with confounders such as socioeconomic status and health care access, contributing to spurious associations [5,6]. Lower participation rates among disadvantaged groups—who bear a disproportionate burden of NCDs—may further distort assessments of health inequalities [7,8]. In parallel, broader technological shifts and declining public engagement with surveys have compounded these challenges [9].

As traditional data sources struggle to capture the complexity and pace of population health dynamics, researchers are increasingly leveraging digital data—such as data from retail transactions, mobile apps, and other web-based activities. This shift reflects the growing momentum of digital epidemiology, which shares the foundational goals of traditional epidemiology—understanding and improving

health at the population level—but sets itself apart by using data not originally generated for epidemiological purposes [10]. One such data resource is supermarket loyalty card data, offering objective, high-frequency insights into population health behaviors.

## Loyalty Card Data: A Tool for Epidemiological Research

Among the vast streams of data generated daily, supermarket loyalty card data—purchase records collected at checkout—provide granular, population-level insights into consumption behaviors relevant to public health [11]. Capturing details on purchased products, transaction frequencies, and store locations, loyalty card data can serve as a cost-effective research tool, as retailers already collect these data. Originally developed for retail marketing, these data are now being repurposed for public health research under strict General Data Protection Regulation (GDPR) guidelines, ensuring deidentified data access and confidentiality [12].

Loyalty card data have several advantages for epidemiological research. First, it provides an objective record of purchases, reducing recall errors and social desirability bias common in self-reported surveys. Second, with widespread adoption across multiple countries [13,14], loyalty cards grant access to large and diverse population samples. Third, as purchase histories—spanning food, health products, and other nondurable goods—are continuously recorded, often across decades, these data enable longitudinal tracking of short- and long-term trends, seasonal variations, and emerging health patterns with a level of granularity and temporal precision rarely achievable through conventional epidemiological data sources.

Demographic metadata from loyalty card programs allows researchers to link purchasing data with existing health datasets (eg, electronic health records, longitudinal surveys) [15,16]. This is particularly useful for studying lifestyle-related risk factors that are difficult to assess via self-reporting,

such as moderate tobacco use or seasonal alcohol consumption. Geospatial data—such as store locations and customer postcodes—support spatial analyses of risky behaviors [17], helping identify regional disparities in food access, exposure to harmful products, and environmental influences.

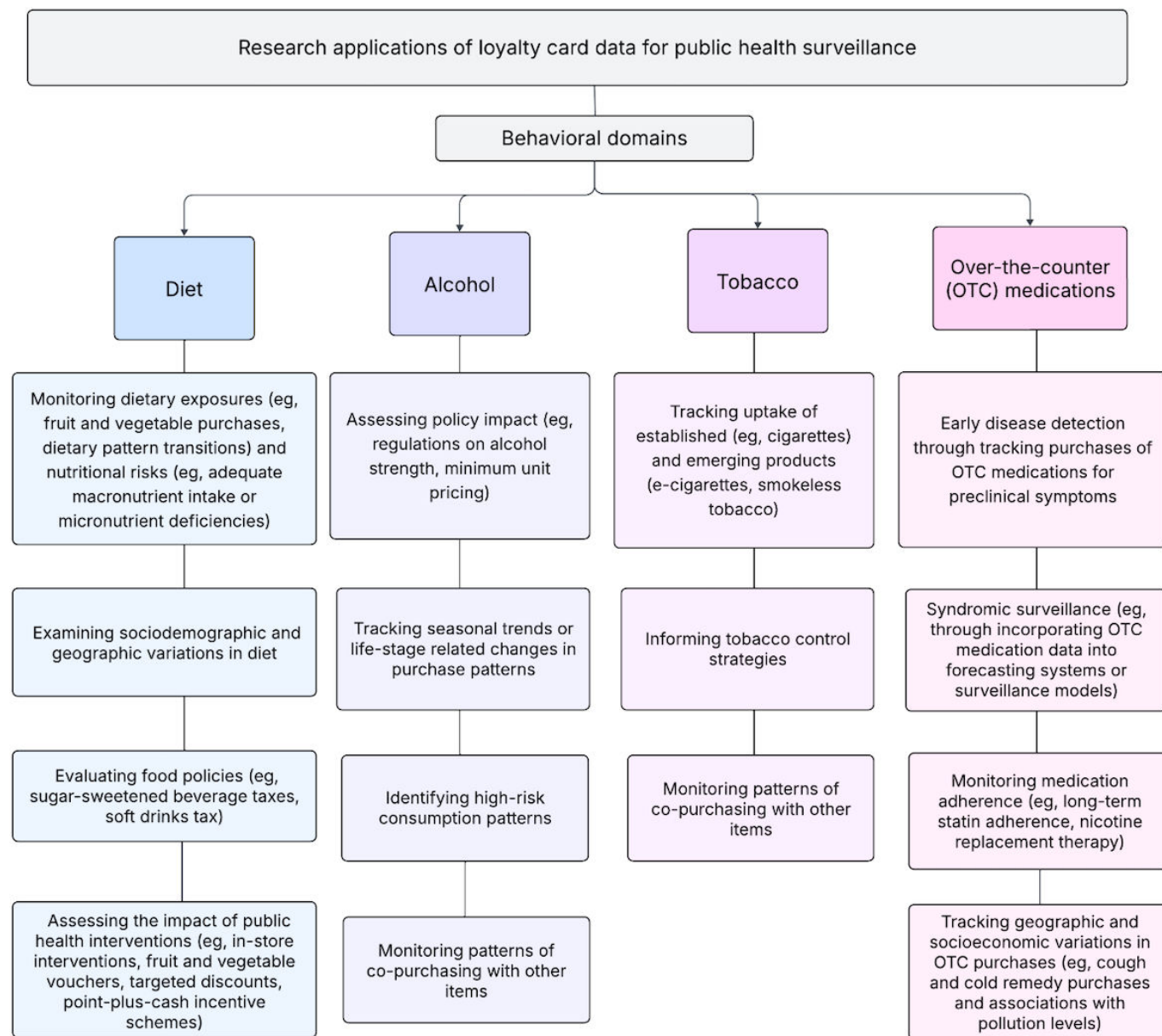
Nonetheless, like any emerging data source, loyalty card data pose methodological challenges that must be addressed [13]. These include privacy concerns, data integration challenges, and methodological constraints. However, advancements in data science and data linkage protocols have enabled successful integration of loyalty card data in large-scale cohort studies such as Avon Longitudinal Study of Parents and Children, showing feasibility for broader implementation [16]. Such approaches are currently most applicable in high-income settings, where supermarkets dominate food supply chains and digital retail is well established [14,18]. As organized retail and mobile-linked loyalty schemes grow in low- and middle-income countries, analogous transaction data may become a scalable option for health surveillance.

## Aims

This paper explores the potential of supermarket loyalty card data as a viable data source for epidemiological surveillance, providing granular, high-frequency insights into population health behaviors. It highlights applications of loyalty card data across 4 key behavioral domains—diet, alcohol, tobacco, and over-the-counter (OTC) medications—and discusses key methodological considerations for leveraging these data in epidemiological research.

## Epidemiological Use Cases Across Key Behavioral Domains

Figure 1 provides an overview of some of the ways in which loyalty card data can be used to provide public health insights across diet, alcohol, tobacco, and OTC medicines.

**Figure 1.** Research applications of loyalty card data across 4 behavioral risk domains.

## Diet

Diet is notoriously difficult to measure due to variations in both the types and quantities of food consumed. As a result, no single “gold-standard” method can address all diet-related questions [14]. Self-reported dietary assessments—such as food frequency questionnaires and 24-hour recall diaries—are limited by recall bias, underreporting, and logistical constraints (eg, limited diet-related questions in multitopic longitudinal studies) [19]. Given the complexity of diet and its outsized impact on health, supplementing self-reports is essential to address gaps in traditional dietary assessments.

Supermarket loyalty card data provide an objective, high-resolution record of food purchases [14] and can serve

as a complementary resource for addressing key epidemiological questions: How do dietary habits and associated health outcomes vary by demographic and socioeconomic factors? How can emerging dietary risks or nutritional deficiencies be detected in real time? Are long-term dietary shifts associated with changes in the consumption of specific nutrients, or preferences for products with specific processing methods or ingredients? Table 1 illustrates the range of diet-related questions that supermarket loyalty card data can be used to address—from mapping sociodemographic differences in food purchasing, through validating purchase-based nutrient estimates, to evaluating fiscal or in-store interventions. These applications are discussed in greater detail in the sections that follow.

**Table 1.** Applications of loyalty card data in studying diet and nutritional intake.

| Research themes and their applications   | Research studies using loyalty card data   |
|--|--|
| Dietary exposures  |  |
| Sociodemographic differences in dietary patterns                                     | Identified dietary patterns using UK supermarket loyalty card data, linked to nutrient intake and socioeconomic characteristics [20].  |
| Regional variation in produce purchases  | Analyzed geographic variation in fruit and vegetable purchases using loyalty card data [21].   |
| Purchase preferences of specific food groups   | Identified clusters with different preferences in protein sources, including transitions away from red meat [22].  |
| Longitudinal trends in purchasing patterns by sociodemographic characteristics       | Modeled trends in produce purchasing over time, stratified by age and income groups [23].  |
| Emerging micronutrient deficiencies  | Detected population-level iodine deficiency risks linked to reduced dairy and increased plant-based milk purchases [24].   |
| Comparison with self-reported data   |  |
| Validation of food purchases and survey data   | Compared loyalty card–based protein purchases with self-reported intake among older adults to assess concordance [25].   |
| Comparison of loyalty card data with self-reported consumption                       | Linked grocery purchases with food frequency questionnaire responses to examine variations across population subgroups [26].   |
| Intervention effectiveness   |  |
| Assessing impact of voucher programs and financial incentives on consumption choices | 1. Evaluated the impact of vouchers on fruit and vegetable purchases in low-income households [27].<br>2. Evaluated a points-plus-cash program promoting fruit and vegetable purchases [28]<br>3. Investigated whether supermarket-based nudges and pricing changes affected ultraprocessed food purchases [29].                                   |
| Impact of nutritional information interventions on food choice                       | 1. Evaluated long-term effects of grocery store podcasts on omega-3 purchases [30].<br>2. Assessed the impact of point-of-sale nutrition information on food selection [31].   |
| Associations with health outcomes  |  |
| Effects of supermarket discounts on health outcomes                                  | 1. Studied effects of discounts on fruits, vegetables, and noncaloric beverages on weight loss and diet quality [32].<br>2. Studied the real-world effects of supermarket nudging and pricing strategies and mobile physical activity coaching on diet quality, food-purchasing behavior, walking behavior, and cardiometabolic risk markers [33]. |
| Food purchases and disease outcomes  | Examined associations between loyalty card–based food purchases and health outcomes including hypertension, high cholesterol, and diabetes [34].   |
| Policy impact assessment   |  |
| Sugar-sweetened beverage tax   | Studied the impact of sugar-sweetened beverage tax effects on beverage purchases, with attention to subgroup-specific effects [35].  |
| UK soft drinks sugar tax   | Assessed impact of UK sugar tax on purchasing patterns [36].   |
| Impact of events and economic shifts   |  |
| COVID-19 impact on fish and seafood purchases  | Explored changes in fish, seafood, and related product purchases during COVID-19 lockdown [37].  |
| Great Recession impact on food purchases   | Analyzed effects of the Great Recession on UK food purchases [38].   |

Assessing Diet and Nutritional Inequalities

One major advantage of loyalty card data is its ability to reveal dietary disparities that traditional surveys often miss. For example, a study linking supermarket loyalty card data with a nutrient composition database found that lower-income households tended to purchase foods lower in fiber, a key marker of diet quality [20]. Similarly, spatial analyses have shown that fruit and vegetable purchases are more common

in affluent areas and among older populations [21]. Such insights can inform targeted nutritional interventions.

Integrating food purchase records with health data and nutritional databases further expands the epidemiological applications of loyalty card data. To give one example, a study linking loyalty card data with medical prescription records found nutrient diversity and caloric intake to be the strongest predictors of “metabolic syndrome” diseases:

hypertension, high cholesterol, and diabetes [34]. By enabling dynamic tracking of diet-disease relationships, these linkages overcome the limitations of survey-based dietary assessments, which provide only fragmented snapshots of exposure. Future research could explore how medical diagnoses, economic conditions, or life transitions shape food choices and associated health outcomes.

These data can also help identify population-level nutritional risks emerging from demographic and cultural shifts, providing early warnings for public health practitioners. For example, a loyalty card data study observed an association between transitioning to nondairy milk alternatives and iodine deficiency risk [24], while another revealed that older adults (aged 55+) in the United Kingdom were less likely to meet recommended levels of protein intake [25]. Importantly, loyalty card data have been shown to be a resource-efficient and moderately valid measure of dietary intake in large samples, reinforcing its value as a complementary tool to self-report surveys [26].

## Policy and Intervention Evaluations

Loyalty card data serve as a powerful tool for evaluating public health policies and interventions. For example, they have been applied to assess the impact of fiscal measures—such as sugar-sweetened beverage taxes and the soft drinks sugar tax—on purchases of targeted items [35,36]. In addition, studies using loyalty card data have examined the effects of economic incentives such as fruit and vegetable vouchers [27], discounts [32], and points-plus-cash programs promoting healthier diets [28], as well as supermarket nudging strategies aimed at reducing ultraprocessed food consumption [29]. By tracking actual purchases of food items rather than relying on self-reported compliance, these data can facilitate more timely and objective evaluations of intervention or policy impact.

## Alcohol

Alcohol epidemiology faces challenges from falling survey response rates and inherent limits of traditional data sources [39]. Self-reported surveys are prone to recall bias and social desirability effects, leading to underreporting—particularly among heavy drinkers—while aggregate sales data lack the granularity to capture individual-level behaviors and demographic differences [40,41].

Loyalty card data provide a cost-effective supplement to traditional surveys by providing detailed, longitudinal records of alcohol purchases—including product type, quantity, and price. These passively collected data reduce respondent burden and self-reporting biases, allowing researchers to address key epidemiological questions: How do purchasing patterns vary by demographic and socioeconomic status? How can we identify seasonal, policy-driven, or life event-related shifts in alcohol consumption? How are these patterns associated with co-occurring risk behaviors and health outcomes?

A notable benefit of loyalty card data is the ability to continuously track alcohol purchase patterns over time, capturing shifts driven by seasonal variations [42], societal

trends (eg, rising abstinence and declining alcohol use) [43] or policy changes [44]. For instance, a Finnish study used loyalty card data to evaluate the impact of a legislative reform allowing grocery stores to sell stronger alcoholic beverages, revealing distinct demographic differences in alcohol purchasing patterns [44]. Similar approaches could be used to assess the impact of policies (eg, minimum unit pricing) or abstinence campaigns (eg, Dry January or Febfast) to determine whether reductions in alcohol purchases persist or relapse over time. By continuously tracking purchases, these insights can enable epidemiologists to assess both immediate policy effects and longer-term behavioral shifts.

Loyalty card data also advance exposure modeling by capturing individual-level alcohol purchasing behaviors over time. A Finnish study demonstrated that beer purchase frequency aligned closely with self-reported beer drinking frequency, supporting the validity of purchase-based estimates of alcohol consumption [45]. Moreover, these data identify co-occurring risk behaviors often overlooked in analyses examining associations between alcohol consumption and health outcomes. This is illustrated by a Finnish study which showed that alcohol purchases frequently coincided with tobacco and unhealthy food purchases [46], while a French study revealed that wine buyers were more likely to purchase healthier foods, whereas beer consumers favored processed and high-fat items [47].

When linked with health records, loyalty card data can serve as a valuable tool to help identify purchasing patterns associated with at-risk populations. These insights can inform adaptive, evidence-based public health strategies, allowing interventions to be responsive to emerging trends in alcohol use and associated risks.

## Tobacco

Epidemiological surveys of tobacco consumption have traditionally focused on cigarette smoking, often overlooking the growing use of less regulated products, such as e-cigarettes, waterpipes, and smokeless tobacco [48]. This narrow scope creates blind spots in understanding the adoption, user demographics, and evolving consumption patterns of newer tobacco products, especially relevant amid the surge in e-cigarette use among young people [49]. Given the fast-changing landscape of tobacco products, more adaptive, real-time surveillance systems are urgently needed.

Loyalty card data, although underused in tobacco research [50], holds valuable information on category preferences, product characteristics (eg, nicotine strength), and purchasing frequency. These data can help address questions like: How are emerging tobacco products, including disposable e-cigarettes, being adopted across different age groups, socioeconomic strata, and geographical areas? How do purchasing behaviors respond to product innovations, industry marketing strategies, or local policy changes—and how are these shifts associated with downstream health outcomes or risk trajectories? To what extent do tobacco purchases co-occur with other risky behaviors—such as alcohol use or unhealthy food consumption—and how do these patterns vary across populations?

In addition, loyalty card data can reveal disparities in tobacco use—especially among marginalized communities targeted by industry marketing—and thereby inform regulatory decisions and public health strategies aimed at curbing emerging tobacco trends and associated risks [2].

## OTC Medications

Loyalty card data from supermarkets and pharmacies provide a continuous, high-frequency record of OTC medication purchases and can enable near real-time monitoring of self-medication behaviors, seasonal illness trends, and potential public health risks.

At the population level, loyalty card data can support passive syndromic surveillance by tracking OTC medication purchases. For example, spikes in purchases of pain relief, gastrointestinal, and respiratory treatments have been shown to predict seasonal illnesses, including influenza, up to 4 weeks in advance—outperforming traditional surveillance models [51]. Integrating OTC medication data into forecasting systems has also improved the accuracy of weekly forecasts of respiratory deaths in England, compared to models using only sociodemographic or weather variables [52]. In addition, these data reveal geographic and socioeconomic differences in OTC medication use. In England, cough and cold remedy purchases were higher in areas with greater exposure to air pollution (PM10 and NO2) [17], while analyses of pain relief, allergy treatments, and sun care products revealed notable geographic and income-based patterns [11]. When linked to health records, loyalty card data could further enhance early warning systems and strengthen disease forecasting models.

OTC medication purchase patterns may also serve as early indicators of disease, helping detect emerging preclinical health changes. Consider a case-control study which found that individuals diagnosed with ovarian cancer showed increased purchases of pain relief and indigestion remedies months before diagnosis, reflecting early, nonspecific symptoms [53]. Such findings suggest that loyalty card data could support early disease detection and inform targeted screening strategies.

Moreover, loyalty card data offer promising avenues for studying medication adherence and treatment continuity. For example, a loyalty card analysis found that frequent store visits and higher spending per trip were associated with better long-term statin adherence [54], while analyses of nicotine replacement therapy purchases revealed patterns of premature discontinuation, mirroring survey findings on cessation challenges [55]. By integrating prescription and loyalty card data, researchers may be able to refine interventions to improve adherence and treatment outcomes.

## Methodological Caveats

While loyalty card data hold promise for epidemiological research, a number of caveats must be addressed. Building on

detailed discussions elsewhere [12,15], we group these into 3 themes: issues intrinsic to loyalty card data, considerations pertinent to epidemiologists, and challenges that arise when integrating with other data sources.

First, because loyalty card data are inherently personal—even when deidentified—robust data-sharing agreements are critical to avoid breaches and maintain public trust [9,34]. This raises broader ethical concerns around the secondary use of commercially collected data for public health research. Informed consent is rarely sought at the point of purchase, raising issues of transparency and autonomy. Data ownership remains ambiguous: while individuals generate the data, retailers control access and use, complicating accountability and benefit-sharing. Commercial partnerships may also create conflicts of interest if contractual terms influence data access, interpretation, or dissemination. Although a detailed discussion of ethical considerations around loyalty card data is beyond the scope of this paper, we encourage readers to consult papers addressing these concerns in detail [12,16,56,57].

From an epidemiological perspective, loyalty card datasets record purchasing—product category, price, quantity, timestamp, and so on—rather than consumption (see the “Variables captured in loyalty card data” column of Table 2 for specific fields available across behavioral domain). Because purchases do not always translate into intake, exposure misclassification is possible [18]. Table 2 summarizes the chief ways this misclassification can arise in each behavioral domain. In dietary assessments, discrepancies may emerge due to household sharing, food waste, and meals consumed outside the home [17]. Similarly, many alcohol purchases occur at off-licences or discount retailers that do not use loyalty programs, and tobacco sales may be underrepresented if products are bought at separate counters or are not linked to loyalty cards. In the case of OTC medications, regulatory changes—such as limits on stimulant laxatives or painkillers—can modify purchasing patterns independently of health needs, complicating trend analyses. Because of these gaps, loyalty card indicators require validation against established measures. Recent studies show this is feasible: LoCard [26] and STRIDE (Supermarket Transaction Records In Dietary Evaluation) [58] studies compared grocery purchases with food frequency questionnaires; a Switzerland-based study calibrated purchase-based diet scores against self-reported food intake [59]; the Supreme Nudge trial compared loyalty card data and self-reported diets in an intervention setting [60]; and a Finnish study found that beer purchase frequency estimated drinking frequency with fair to good accuracy, depending on how much participants used the same retailer [45]. Although purchases cannot reflect consumption perfectly, these validation studies demonstrate that well-chosen indicators can produce meaningful behavioral proxies. Future work should focus on identifying which product-level or category-level variables track well-established risk exposures most reliably, thereby reducing bias in population surveillance.

**Table 2.** Summary of research applications and methodological limitations of loyalty card data across 4 behavioral risk domains.

| Behavioral domain            | Variables captured in loyalty card data   | Key methodological limitations  |
|------------------------------|---|---|
| Diet                         | Types of food purchased (eg, fresh produce and processed foods), quantity, frequency of purchase, nutrient profile via linkage with food databases  | Does not capture actual consumption; household-level purchases may not reflect individual intake; limited coverage of food wastage or food eaten at out-of-home sectors; often requires linkages to other databases to extract information on ingredients or nutritional content. |
| Alcohol                      | Product type (beer, wine, and spirits), quantity, timing, and price sensitivity   | Under coverage of alcohol bought at off-licences or other retailers; purchases may not reflect consumption; limited product-level detail on alcohol content (ABV).  |
| Tobacco                      | Product category (eg, cigarette, e-cigarette, or smokeless tobacco); product brand; product characteristics (eg, nicotine strength), price, purchase frequency, and quantity                                    | Low volume of purchases captured due to separate counters or exclusion from loyalty schemes; may lack user demographics; regulatory changes may skew trends.  |
| Over-the-counter medications | Product category (eg, pain relief medications, cough and cold remedies, allergy medicines, period products, laxatives, nicotine replacement therapy products, and supplements), quantity, frequency, and timing | Purchases may not indicate use; regulatory limits (eg, on stimulant laxatives or painkillers) can alter purchasing patterns independently of health needs.  |

Representativeness is another core limitation: not all consumers participate in loyalty programs, and some may pay with nontraceable methods or shop across multiple retailers, introducing selection bias [26]. Supermarket loyalty card data may overrepresent certain demographic groups, including older adults [20,21], females [60], smaller households [61], or those with higher education [21,22] and income [28,61] compared to the general population. This may lead to underrepresentation of vulnerable groups—failure to account for which could potentially skew behavioral estimates and mask or misrepresent the burden of risk among socioeconomically disadvantaged populations—compromising the validity of inferences about health inequalities. Future studies should consider weighting strategies [13], linkage to existing population datasets [16], or validation against population-representative samples to assess and correct for these biases [13].

Moreover, since loyalty card datasets are commercially produced, they may suffer from missing entries, inconsistent categorization, or insufficient detail in product groupings [17]. For epidemiologists, distinguishing between “risky” and “non-risky” consumption can become complicated when product categories are not clearly defined or standardized. Moreover, changes in retailer product categorizations and loyalty program structures can affect data consistency over time, complicating longitudinal analyses and necessitating regular model recalibration. Finally, causal inference may be particularly challenging: reverse causation can arise if individuals alter their purchasing patterns after a health diagnosis, while unmeasured confounders—such as local food environments or targeted marketing—can produce spurious associations [32]. To strengthen scope for causal interpretations, researchers must design studies with clearly defined hypotheses, reliable exposure indicators, and

statistical methods that account for temporal variations and potential confounders [51].

From a technical standpoint, processing large-scale loyalty card data requires sophisticated preprocessing, substantial computational power, and statistical expertise to extract meaningful patterns while avoiding overfitting [62]. Advanced machine learning and statistical techniques are crucial for reconciling data inconsistencies, detecting outliers, and managing large-scale transactional datasets [62]; however, these methods often exceed the routine methodological toolkit of many epidemiologists, highlighting the need for interdisciplinary collaboration with data scientists and health informatics specialists [21]. By drawing on established protocols designed for loyalty card linkages with existing cohort studies [16], maintaining transparent governance structures, and fostering multisector collaborations, researchers can better leverage loyalty card data to generate insights into population health.

Importantly, realizing the potential of loyalty card data for epidemiological research requires integrating these data with existing cohort studies and health records. Such linkages demand careful handling to preserve privacy, ensure accurate matching, and mitigate risks of reidentification [14]. Inaccurate linkage can lead to spurious associations, while reliance on single-retailer datasets may provide only a partial view of health behaviors. Although linking data across retailers could provide richer insights, ensuring compatibility among disparate data systems and navigating strict data privacy laws remain significant hurdles [26].

Beyond loyalty card data integration with cohort studies and health records, future research should explore linking loyalty card data with emerging digital health sources. These include smartphone health apps (eg, diet tracking or

menstrual cycle apps), wearable devices (eg, fitness trackers and sleep monitors), and even social media platforms (eg, posts or interactions reflecting mood or stress). Such linkages would enable more dynamic and multidimensional behavioral phenotyping—capturing not only what people buy but also when, why, and how health behaviors cluster in real time.

## Implications for Policy Makers and Public Health Agencies

Beyond academic research, loyalty card data can be embedded directly into routine public health intelligence systems. Because loyalty card data are generated daily, public health agencies may be able to construct near-real-time dashboards that flag unusual spikes in purchases of sentinel products—for example, analgesics or cough-and-cold remedies—as early warnings of respiratory outbreaks weeks before clinical reports become available [51,52]. At finer spatial scales, anonymized postcode-level purchase maps can reveal “hot spots” of risky behaviors (eg, clusters of high-strength alcohol purchases or disposable e-cigarette sales) that could enable local authorities to deploy targeted place-based health-promotion teams or licensing controls. Finally, because loyalty card data streams document population response to fiscal or regulatory actions (such as sugar-sweetened beverage taxes or alcohol price changes), policymakers can build rapid feedback loops that track equity-stratified effects within weeks rather than years and iteratively refine policies while they are still politically and economically tractable.

## Acknowledgments

This work was supported by UKRI Future Leaders Fellowship (MR/T043520/1) and ESRC Smart Data Accelerator Award (ES/Y010973/1) to Anya Skatova.

## Authors' Contributions

Alisha Suhag drafted the manuscript. Alisha Suhag and RB conducted the literature review. Anya Skatova and Alisha Suhag conceptualized the paper. RB and Anya Skatova revised it critically for important intellectual content.

## Conflicts of Interest

None declared.

## References

1. Mayor S. Non-communicable diseases now cause two thirds of deaths worldwide. *BMJ*. 2016;355:i5456. [doi: [10.1136/bmj.i5456](https://doi.org/10.1136/bmj.i5456)]
2. Extending Life: Progress and Achievements in 2017 of the WHO European Office for the Prevention and Control of Noncommunicable Diseases. World Health Organization. Regional Office for Europe.; 2017. URL: <https://iris.who.int/handle/10665/343965> [Accessed 2025-07-24]
3. Taylor AE, Jones HJ, Sallis H, et al. Exploring the association of genetic factors with participation in the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. Aug 1, 2018;47(4):1207-1216. [doi: [10.1093/ije/dyy060](https://doi.org/10.1093/ije/dyy060)] [Medline: [29800128](https://pubmed.ncbi.nlm.nih.gov/29800128/)]
4. Kilian C, Manthey J, Probst C, et al. Why is per capita consumption underestimated in alcohol surveys? Results from 39 surveys in 23 European countries. *Alcohol Alcohol*. Aug 14, 2020;55(5):554-563. [doi: [10.1093/alcalc/agaa048](https://doi.org/10.1093/alcalc/agaa048)] [Medline: [32491170](https://pubmed.ncbi.nlm.nih.gov/32491170/)]
5. Rinsky JL, Richardson DB, Wing S, et al. Assessing the potential for bias from nonresponse to a study follow-up interview: an example from the Agricultural Health Study. *Am J Epidemiol*. Aug 15, 2017;186(4):395-404. [doi: [10.1093/aje/kwx098](https://doi.org/10.1093/aje/kwx098)] [Medline: [28486574](https://pubmed.ncbi.nlm.nih.gov/28486574/)]

## Conclusions

Loyalty card data represent a transformative frontier for public health surveillance. By providing unprecedented granularity and scale in capturing real-world behaviors—diet, alcohol, tobacco, and OTC medication—these data can not only supplement but, in many cases, substantially enhance insights gained by traditional self-report methods. When linked with existing health records, loyalty card data have the potential to improve exposure measurement, enable dynamic risk modeling, and accelerate the evaluation of policies and public health interventions. Unlocking its full potential will demand bold investment in technical capacity, stronger data governance, and cross-sector collaboration. But the payoff is clear—a more agile and data-rich public health ecosystem capable of responding to emerging threats with greater speed and precision.

## Ethical Considerations

This viewpoint does not report new research involving human participants, animals, or identifiable personal data. All examples and datasets referenced are drawn from previously published sources already in the public domain; therefore, formal ethics committee review and informed consent were not required.

6. Keiding N, Louis TA. Perils and potentials of self-selected entry to epidemiological studies and surveys. *J R Stat Soc Ser A Stat Soc*. Feb 1, 2016;179(2):319-376. [doi: [10.1111/rssa.12136](https://doi.org/10.1111/rssa.12136)]
7. Lorant V, Demarest S, Miermans PJ, Van Oyen H. Survey error in measuring socio-economic risk factors of health status: a comparison of a survey and a census. *Int J Epidemiol*. Dec 2007;36(6):1292-1299. [doi: [10.1093/ije/dym191](https://doi.org/10.1093/ije/dym191)] [Medline: [17898025](https://pubmed.ncbi.nlm.nih.gov/17898025/)]
8. Enzenbach C, Wicklein B, Wirkner K, Loeffler M. Evaluating selection bias in a population-based cohort study with low baseline participation: the LIFE-Adult-Study. *BMC Med Res Methodol*. Jul 1, 2019;19(1):135. [doi: [10.1186/s12874-019-0779-8](https://doi.org/10.1186/s12874-019-0779-8)] [Medline: [31262266](https://pubmed.ncbi.nlm.nih.gov/31262266/)]
9. Liber AC, Warner KE. Has underreporting of cigarette consumption changed over time? Estimates derived from US National Health Surveillance Systems between 1965 and 2015. *Am J Epidemiol*. Jan 1, 2018;187(1):113-119. [doi: [10.1093/aje/kwx196](https://doi.org/10.1093/aje/kwx196)] [Medline: [28595348](https://pubmed.ncbi.nlm.nih.gov/28595348/)]
10. Salathé M. Digital epidemiology: what is it, and where is it going? *Life Sci Soc Policy*. Jan 4, 2018;14(1):1. [doi: [10.1186/s40504-017-0065-7](https://doi.org/10.1186/s40504-017-0065-7)] [Medline: [29302758](https://pubmed.ncbi.nlm.nih.gov/29302758/)]
11. Davies A, Green MA, Riddlesden D, Singleton AD. Using loyalty card records and machine learning to understand how self-medication purchasing behaviours vary seasonally in England, 2012–2014. *Appl Mark Anal*. 2020;5(4):354-370. [doi: [10.69554/MJDF4395](https://doi.org/10.69554/MJDF4395)]
12. Skatova A. Overcoming biases of individual level shopping history data in health research. *NPJ Digit Med*. Sep 30, 2024;7(1):264. [doi: [10.1038/s41746-024-01231-4](https://doi.org/10.1038/s41746-024-01231-4)] [Medline: [39349949](https://pubmed.ncbi.nlm.nih.gov/39349949/)]
13. Vuorinen AL, Erkkola M, Fogelholm M, et al. Characterization and correction of bias due to nonparticipation and the degree of loyalty in large-scale Finnish loyalty card data on grocery purchases: cohort study. *J Med Internet Res*. Jul 15, 2020;22(7):e18059. [doi: [10.2196/18059](https://doi.org/10.2196/18059)] [Medline: [32459633](https://pubmed.ncbi.nlm.nih.gov/32459633/)]
14. Jenneson VL, Pontin F, Greenwood DC, Clarke GP, Morris MA. A systematic review of supermarket automated electronic sales data for population dietary surveillance. *Nutr Rev*. May 9, 2022;80(6):1711-1722. [doi: [10.1093/nutrit/nuab089](https://doi.org/10.1093/nutrit/nuab089)] [Medline: [34757399](https://pubmed.ncbi.nlm.nih.gov/34757399/)]
15. Nevalainen J, Erkkola M, Saarijärvi H, Näppilä T, Fogelholm M. Large-scale loyalty card data in health research. *Digit Health*. 2018;4:2055207618816898. [doi: [10.1177/2055207618816898](https://doi.org/10.1177/2055207618816898)] [Medline: [30546912](https://pubmed.ncbi.nlm.nih.gov/30546912/)]
16. Skatova A, Boyd A. A protocol for linking participants' retailer 'loyalty card' records into the Avon Longitudinal Study of Parents and Children (ALSPAC). *Wellcome Open Res*. 2024;8:99. [doi: [10.12688/wellcomeopenres.18900.2](https://doi.org/10.12688/wellcomeopenres.18900.2)]
17. Davies A, Green MA, Singleton AD. Using machine learning to investigate self-medication purchasing in England via high street retailer loyalty card data. *PLoS One*. 2018;13(11):e0207523. [doi: [10.1371/journal.pone.0207523](https://doi.org/10.1371/journal.pone.0207523)] [Medline: [30452481](https://pubmed.ncbi.nlm.nih.gov/30452481/)]
18. Tin ST, Mhurchu CN, Bullen C. Supermarket sales data: feasibility and applicability in population food and nutrition monitoring. *Nutr Rev*. Jan 2007;65(1):20-30. [doi: [10.1111/j.1753-4887.2007.tb00264.x](https://doi.org/10.1111/j.1753-4887.2007.tb00264.x)] [Medline: [17310856](https://pubmed.ncbi.nlm.nih.gov/17310856/)]
19. Ravelli MN, Schoeller DA. Traditional self-reported dietary instruments are prone to inaccuracies and new approaches are needed. *Front Nutr*. 2020;7:90. [doi: [10.3389/fnut.2020.00090](https://doi.org/10.3389/fnut.2020.00090)] [Medline: [32719809](https://pubmed.ncbi.nlm.nih.gov/32719809/)]
20. Clark SD, Shute B, Jenneson V, Rains T, Birkin M, Morris MA. Dietary patterns derived from UK supermarket transaction data with nutrient and socioeconomic profiles. *Nutrients*. Apr 27, 2021;13(5):1481. [doi: [10.3390/nu13051481](https://doi.org/10.3390/nu13051481)] [Medline: [33925712](https://pubmed.ncbi.nlm.nih.gov/33925712/)]
21. Jenneson V, Clarke GP, Greenwood DC, et al. Exploring the geographic variation in fruit and vegetable purchasing behaviour using supermarket transaction data. *Nutrients*. Dec 30, 2021;14(1):177. [doi: [10.3390/nu14010177](https://doi.org/10.3390/nu14010177)] [Medline: [35011053](https://pubmed.ncbi.nlm.nih.gov/35011053/)]
22. Erkkola M, Kinnunen SM, Vepsäläinen HR, et al. A slow road from meat dominance to more sustainable diets: an analysis of purchase preferences among Finnish loyalty-card holders. *PLOS Sustain Transform*. 2022;1(6):e0000015. [doi: [10.1371/journal.pstr.0000015](https://doi.org/10.1371/journal.pstr.0000015)]
23. Fernandez ID, Johnson BA, Wixom N, et al. Longitudinal trends in produce purchasing behavior: a descriptive study of transaction level data from loyalty card households. *Nutr J*. Nov 8, 2022;21(1):67. [doi: [10.1186/s12937-022-00814-9](https://doi.org/10.1186/s12937-022-00814-9)] [Medline: [36348423](https://pubmed.ncbi.nlm.nih.gov/36348423/)]
24. Mansilla R, Long G, Welham S, et al. Detecting iodine deficiency risks from dietary transitions using shopping data. *Sci Rep*. Jan 10, 2024;14(1):1017. [doi: [10.1038/s41598-023-50180-7](https://doi.org/10.1038/s41598-023-50180-7)] [Medline: [38200032](https://pubmed.ncbi.nlm.nih.gov/38200032/)]
25. Green MA, Watson AW, Brunstrom JM, et al. Comparing supermarket loyalty card data with traditional diet survey data for understanding how protein is purchased and consumed in older adults for the UK, 2014-16. *Nutr J*. Aug 13, 2020;19(1):83. [doi: [10.1186/s12937-020-00602-3](https://doi.org/10.1186/s12937-020-00602-3)] [Medline: [32791968](https://pubmed.ncbi.nlm.nih.gov/32791968/)]
26. Vepsäläinen H, Nevalainen J, Kinnunen S, et al. Do we eat what we buy? Relative validity of grocery purchase data as an indicator of food consumption in the LoCard study. *Br J Nutr*. Nov 14, 2022;128(9):1780-1788. [doi: [10.1017/S0007114521004177](https://doi.org/10.1017/S0007114521004177)] [Medline: [34657639](https://pubmed.ncbi.nlm.nih.gov/34657639/)]

27. Thomas M, Moore JB, Onuselogu DA, et al. Supermarket top-up of Healthy Start vouchers increases fruit and vegetable purchases in low-income households. *Nutr Bull.* Sep 2023;48(3):353-364. [doi: [10.1111/nbu.12627](https://doi.org/10.1111/nbu.12627)] [Medline: [37501220](https://pubmed.ncbi.nlm.nih.gov/37501220/)]
28. Panzone LA, Tocco B, Brečić R, Gorton M. Healthy foods, healthy sales? Cross-category effects of a loyalty program promoting sales of fruit and vegetables. *J Retail.* Mar 2024;100(1):85-103. [doi: [10.1016/j.jretai.2023.12.002](https://doi.org/10.1016/j.jretai.2023.12.002)]
29. Mackenbach JD, Pinho MGM, Stuber JM, et al. The effects of nudging and pricing strategies on the availability and purchases of ultra-processed foods: a secondary analysis of the Supreme Nudge trial. *Appetite.* Oct 1, 2024;201:107599. [doi: [10.1016/j.appet.2024.107599](https://doi.org/10.1016/j.appet.2024.107599)] [Medline: [38992569](https://pubmed.ncbi.nlm.nih.gov/38992569/)]
30. Bangia D, Shaffner DW, Palmer-Keenan DM. A point-of-purchase intervention using grocery store tour podcasts about omega-3s increases long-term purchases of omega-3-rich food items. *J Nutr Educ Behav.* Jun 2017;49(6):475-480. [doi: [10.1016/j.jneb.2017.02.008](https://doi.org/10.1016/j.jneb.2017.02.008)] [Medline: [28377095](https://pubmed.ncbi.nlm.nih.gov/28377095/)]
31. Nikolova HD, Inman JJ. Healthy choice: the effect of simplified point-of-sale nutritional information on consumer food choice behavior. *J Mark Res.* Dec 2015;52(6):817-835. [doi: [10.1509/jmr.13.0270](https://doi.org/10.1509/jmr.13.0270)]
32. Poskute AS, Ang IYH, Rahman N, Geliebter A. Effects of discounting fruits, vegetables, and noncaloric beverages in New York City supermarkets on purchasing, intake, and weight. *Obesity (Silver Spring).* Jul 2024;32(7):1290-1301. [doi: [10.1002/oby.24058](https://doi.org/10.1002/oby.24058)] [Medline: [38932711](https://pubmed.ncbi.nlm.nih.gov/38932711/)]
33. Stuber JM, Mackenbach JD, de Bruijn GJ, et al. Real-world nudging, pricing, and mobile physical activity coaching was insufficient to improve lifestyle behaviours and cardiometabolic health: the Supreme Nudge parallel cluster-randomised controlled supermarket trial. *BMC Med.* Feb 2, 2024;22(1):52. [doi: [10.1186/s12916-024-03268-4](https://doi.org/10.1186/s12916-024-03268-4)] [Medline: [38303069](https://pubmed.ncbi.nlm.nih.gov/38303069/)]
34. Aiello LM, Schifanella R, Quercia D, Del Prete L. Large-scale and high-resolution analysis of food purchases and health outcomes. *EPJ Data Sci.* Dec 2019;8(1):1-22. [doi: [10.1140/epjds/s13688-019-0191-y](https://doi.org/10.1140/epjds/s13688-019-0191-y)]
35. Fichera E, Mora T, Lopez-Valcarcel BG, Roche D. How do consumers respond to “sin taxes”? New evidence from a tax on sugary drinks. *Soc Sci Med.* Apr 2021;274:113799. [doi: [10.1016/j.socscimed.2021.113799](https://doi.org/10.1016/j.socscimed.2021.113799)] [Medline: [33684702](https://pubmed.ncbi.nlm.nih.gov/33684702/)]
36. Fearne A, Borzino N, De La Iglesia B, Moffatt P, Robbins M. Using supermarket loyalty card data to measure the differential impact of the UK soft drink sugar tax on buyer behaviour. *J Agric Econ.* Jun 2022;73(2):321-337. URL: <https://onlinelibrary.wiley.com/toc/14779552/73/2> [doi: [10.1111/1477-9552.12462](https://doi.org/10.1111/1477-9552.12462)]
37. de la Iglesia R, García-González Á, Achón M, Varela-Moreiras G, Alonso Aperte E. Fish, seafood, and fish products purchasing habits in the Spanish population during COVID-19 lockdown. *Int J Environ Res Public Health.* Sep 15, 2022;19(18):11624. [doi: [10.3390/ijerph191811624](https://doi.org/10.3390/ijerph191811624)] [Medline: [36141898](https://pubmed.ncbi.nlm.nih.gov/36141898/)]
38. Griffith R, O’Connell M, Smith K. Shopping around: how households adjusted food spending over the great recession. *Economica.* Apr 2016;83(330):247-280. URL: <https://onlinelibrary.wiley.com/toc/14680335/83/330> [doi: [10.1111/ecca.12166](https://doi.org/10.1111/ecca.12166)]
39. Meiklejohn J, Connor J, Kypri K. The effect of low survey response rates on estimates of alcohol consumption in a general population survey. *PLoS One.* 2012;7(4):e35527. [doi: [10.1371/journal.pone.0035527](https://doi.org/10.1371/journal.pone.0035527)] [Medline: [22532858](https://pubmed.ncbi.nlm.nih.gov/22532858/)]
40. Stockwell T, Zhao J, Greenfield T, Li J, Livingston M, Meng Y. Estimating under- and over-reporting of drinking in national surveys of alcohol consumption: identification of consistent biases across four English-speaking countries. *Addiction.* Jul 2016;111(7):1203-1213. [doi: [10.1111/add.13373](https://doi.org/10.1111/add.13373)] [Medline: [26948693](https://pubmed.ncbi.nlm.nih.gov/26948693/)]
41. Rehm J, Kilian C, Rovira P, Shield KD, Manthey J. The elusiveness of representativeness in general population surveys for alcohol. *Drug Alcohol Rev.* Feb 2021;40(2):161-165. [doi: [10.1111/dar.13148](https://doi.org/10.1111/dar.13148)] [Medline: [32830351](https://pubmed.ncbi.nlm.nih.gov/32830351/)]
42. Autio R, Virta J, Nordhausen K, Fogelholm M, Erkkola M, Nevalainen J. Tensorial principal component analysis in detecting temporal trajectories of purchase patterns in loyalty card data: retrospective cohort study. *J Med Internet Res.* Dec 15, 2023;25:e44599. [doi: [10.2196/44599](https://doi.org/10.2196/44599)] [Medline: [38100168](https://pubmed.ncbi.nlm.nih.gov/38100168/)]
43. Katainen A, Uusitalo L, Saarijärvi H, et al. Who buys non-alcoholic beer in Finland? Sociodemographic characteristics and associations with regular beer purchases. *Int J Drug Policy.* Mar 2023;113:103962. [doi: [10.1016/j.drugpo.2023.103962](https://doi.org/10.1016/j.drugpo.2023.103962)] [Medline: [36746032](https://pubmed.ncbi.nlm.nih.gov/36746032/)]
44. Uusitalo L, Nevalainen J, Rahkonen O, et al. Changes in alcohol purchases from grocery stores after authorising the sale of stronger beverages: the case of the Finnish alcohol legislation reform in 2018. *Nordisk Alkohol Nark.* Dec 2022;39(6):589-604. [doi: [10.1177/14550725221082364](https://doi.org/10.1177/14550725221082364)] [Medline: [36452445](https://pubmed.ncbi.nlm.nih.gov/36452445/)]
45. Lintonen T, Uusitalo L, Erkkola M, et al. Grocery purchase data in the study of alcohol use - a validity study. *Drug Alcohol Depend.* Sep 1, 2020;214:108145. [doi: [10.1016/j.drugalcdep.2020.108145](https://doi.org/10.1016/j.drugalcdep.2020.108145)] [Medline: [32663761](https://pubmed.ncbi.nlm.nih.gov/32663761/)]
46. Uusitalo L, Erkkola M, Lintonen T, Rahkonen O, Nevalainen J. Alcohol expenditure in grocery stores and their associations with tobacco and food expenditures. *BMC Public Health.* Jun 20, 2019;19(1):787. [doi: [10.1186/s12889-019-7096-3](https://doi.org/10.1186/s12889-019-7096-3)] [Medline: [31221122](https://pubmed.ncbi.nlm.nih.gov/31221122/)]
47. Hansel B, Roussel R, Diguët V, Deplaud A, Chapman MJ, Bruckert E. Relationships between consumption of alcoholic beverages and healthy foods: the French supermarket cohort of 196,000 subjects. *Eur J Prev Cardiol.* Feb 2015;22(2):215-222. [doi: [10.1177/2047487313506829](https://doi.org/10.1177/2047487313506829)] [Medline: [24065742](https://pubmed.ncbi.nlm.nih.gov/24065742/)]

48. Theilmann M, Lemp JM, Winkler V, et al. Patterns of tobacco use in low and middle income countries by tobacco product and sociodemographic characteristics: nationally representative survey data from 82 countries. *BMJ*. Aug 30, 2022;378:e067582. [doi: [10.1136/bmj-2021-067582](https://doi.org/10.1136/bmj-2021-067582)] [Medline: [36041745](https://pubmed.ncbi.nlm.nih.gov/36041745/)]
49. Sun J, Xi B, Ma C, Zhao M, Bovet P. Prevalence of e-cigarette use and its associated factors among youths aged 12 to 16 years in 68 countries and territories: Global Youth Tobacco Survey, 2012-2019. *Am J Public Health*. Apr 2022;112(4):650-661. [doi: [10.2105/AJPH.2021.306686](https://doi.org/10.2105/AJPH.2021.306686)] [Medline: [35319939](https://pubmed.ncbi.nlm.nih.gov/35319939/)]
50. Jackson SE, Tattan-Birch H, Brown J. Trends in where people buy their vaping products and differences by user and device characteristics: a population study in England, 2016-23. *Addiction*. Mar 2025;120(3):539-548. [doi: [10.1111/add.16387](https://doi.org/10.1111/add.16387)] [Medline: [37963674](https://pubmed.ncbi.nlm.nih.gov/37963674/)]
51. Miliou I, Xiong X, Rinzivillo S, et al. Predicting seasonal influenza using supermarket retail records. *PLOS Comput Biol*. Jul 2021;17(7):e1009087. [doi: [10.1371/journal.pcbi.1009087](https://doi.org/10.1371/journal.pcbi.1009087)] [Medline: [34252075](https://pubmed.ncbi.nlm.nih.gov/34252075/)]
52. Dolan E, Goulding J, Marshall H, Smith G, Long G, Tata LJ. Assessing the value of integrating national longitudinal shopping data into respiratory disease forecasting models. *Nat Commun*. Nov 21, 2023;14(1):7258. [doi: [10.1038/s41467-023-42776-4](https://doi.org/10.1038/s41467-023-42776-4)] [Medline: [37990023](https://pubmed.ncbi.nlm.nih.gov/37990023/)]
53. Brewer HR, Hirst Y, Chadeau-Hyam M, Johnson E, Sundar S, Flanagan JM. Association between purchase of over-the-counter medications and ovarian cancer diagnosis in the Cancer Loyalty Card Study (CLOCS): Observational Case-Control Study. *JMIR Public Health Surveill*. Jan 26, 2023;9(1):e41762. [doi: [10.2196/41762](https://doi.org/10.2196/41762)] [Medline: [36701184](https://pubmed.ncbi.nlm.nih.gov/36701184/)]
54. Krumme AA, Sanf lix-Gimeno G, Franklin JM, et al. Can purchasing information be used to predict adherence to cardiovascular medications? An analysis of linked retail pharmacy and insurance claims data. *BMJ Open*. Nov 9, 2016;6(11):e011015. [doi: [10.1136/bmjopen-2015-011015](https://doi.org/10.1136/bmjopen-2015-011015)] [Medline: [28186924](https://pubmed.ncbi.nlm.nih.gov/28186924/)]
55. Timberlake DS, Joensuu J, Kurko T, Rimpel  AH, Nevalainen J. Examining retail purchases of cigarettes and nicotine replacement therapy in Finland. *Tob Induc Dis*. 2019;17:39. [doi: [10.18332/tid/108537](https://doi.org/10.18332/tid/108537)] [Medline: [31516482](https://pubmed.ncbi.nlm.nih.gov/31516482/)]
56. Dolan EH, Shiells K, Goulding J, Skatova A. Public attitudes towards sharing loyalty card data for academic health research: a qualitative study. *BMC Med Ethics*. Jun 7, 2022;23(1):58. [doi: [10.1186/s12910-022-00795-8](https://doi.org/10.1186/s12910-022-00795-8)] [Medline: [35672737](https://pubmed.ncbi.nlm.nih.gov/35672737/)]
57. Skatova A, McDonald R, Ma S, Maple C. Unpacking privacy: valuation of personal data protection. *PLoS ONE*. 2023;18(5):e0284581. [doi: [10.1371/journal.pone.0284581](https://doi.org/10.1371/journal.pone.0284581)] [Medline: [37134067](https://pubmed.ncbi.nlm.nih.gov/37134067/)]
58. Jenneson V, Greenwood DC, Clarke GP, et al. Supermarket transaction records In dietary evaluation: the STRIDE study: validation against self-reported dietary intake. *Public Health Nutr*. Dec 2023;26(12):2663-2676. [doi: [10.1017/S1368980023001842](https://doi.org/10.1017/S1368980023001842)] [Medline: [37671553](https://pubmed.ncbi.nlm.nih.gov/37671553/)]
59. Wu J, Fuchs K, Lian J, et al. Estimating dietary intake from grocery shopping data-a comparative validation of relevant indicators in Switzerland. *Nutrients*. Dec 29, 2021;14(1):159. [doi: [10.3390/nu14010159](https://doi.org/10.3390/nu14010159)] [Medline: [35011033](https://pubmed.ncbi.nlm.nih.gov/35011033/)]
60. Colizzi C, Stuber JM, van der Schouw YT, Beulens JWJ. Are food and beverage purchases reflective of dietary intake? Validity of supermarket purchases as indicator of diet quality in the Supreme Nudge Trial. *Br J Nutr*. Nov 28, 2024;132(10):1394-1402. [doi: [10.1017/S0007114524002630](https://doi.org/10.1017/S0007114524002630)] [Medline: [39501842](https://pubmed.ncbi.nlm.nih.gov/39501842/)]
61. Fogelholm M, Veps l inen H, Meinil  J, et al. The dynamics in food selection stemming from price awareness and perceived income adequacy: a cross-sectional study using 1-year loyalty card data. *Am J Clin Nutr*. May 2024;119(5):1346-1353. [doi: [10.1016/j.ajcnut.2024.03.003](https://doi.org/10.1016/j.ajcnut.2024.03.003)] [Medline: [38458401](https://pubmed.ncbi.nlm.nih.gov/38458401/)]
62. Dolan E, Goulding J, Lang AR, Tata LJ, Skatova A. The feasibility of using individual-level loyalty card data for disease surveillance: a pilot study on COVID-19. Preprint posted online on 2024. [doi: [10.21203/rs.3.rs-5239099/v1](https://doi.org/10.21203/rs.3.rs-5239099/v1)]

## Abbreviations

**GDPR:** General Data Protection Regulation

**NCD:** noncommunicable disease

**OTC:** over-the-counter

**STRIDE:** Supermarket Transaction Records In Dietary Evaluation

*Edited by Amaryllis Mavragani; peer-reviewed by Chibuzo Onah, Ravi Teja Potla; submitted 09.04.2025; final revised version received 13.06.2025; accepted 14.06.2025; published 06.08.2025*

*Please cite as:*

Suhag A, Burgess R, Skatova A

*Shopping Data for Population Health Surveillance: Opportunities, Challenges, and Future Directions*

*J Med Internet Res* 2025;27:e75720

URL: <https://www.jmir.org/2025/1/e75720>

doi: [10.2196/75720](https://doi.org/10.2196/75720)

© Alisha Suhag, Romana Burgess, Anya Skatova. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 06.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org>, as well as this copyright and license information must be included.